

Univerzita Karlova  
Filozofická fakulta  
Ústav informačních studií a knihovnictví  
Študijný program: informační studia a knihovnictví

Bc. Beáta Jakubcová

METODIKA HODNOCENÍ V SPOLEČNOSTI GOOGLE  
METHODOLOGY FOR EVALUATION IN GOOGLE COMPANY

Diplomová práce

Praha 2013

Vedúci práce: Ing. Martin Souček, Ph.D.

Podakovanie:

Rada by som poďakovala vedúcemu mojej práce Ing. Martinovi Součkovi za odborné konzultácie a trpezlivosť, ktorú so mnou mal a všetkým expertom z oblasti internetu a SEO, ktorí mi svojím zanietením pomohli zorientovať sa v tejto oblasti.

Zároveň by som chcela vyjadriť vďaku zakladateľom spoločnosti Google Larry Pageovi a Sergeyovi Brinovi, pretože bez nich by táto práca nikdy nemohla vzniknúť.

Prehlásenie:

Prehlasujem, že som diplomovú prácu vykonala samostatne, že som riadne citovala všetky použité pramene a literatúru a že práca nebola využitá v rámci iného vysokoškolského štúdia, alebo na získanie iného alebo rovnakého titulu.

V Prahe dňa 8.4.2013

## **Abstrakt**

Diplomová práca popisuje vyhľadávač spoločnosti Google a jeho spôsob hodnotenia webových stránok, na základe ktorého ich radí pri zobrazení na výsledkovej stránke vyhľadávania. Okrem popisu fungovania vyhľadávačov všeobecne sa práca zameriava na odlišnosti medzi Google a ostatnými vyhľadávačmi a popisuje charakteristiky a princípy navrhnuté jeho zakladateľmi Larrym Pageom a Sergeyom Brinom, ktoré ho robia jedinečným. V ďalšej časti sú načrtnuté niektoré z množstva signálov, ktoré sú vyhodnocované pre každú webovú stránku. Z nich je najväčší priestor venovaný PageRanku ako ťažnej myšlienke hodnotenia stránok na základe ich odkazovej schémy, ktorá vychádza z princípov citačnej analýzy. Hodnotiace faktory sú v práci analyzované na konkrétnych príkladoch a záver práce popisuje zmeny, ktoré technológiu vyhľadávača Google sprevádzali od jeho začiatkov v roku 1998 až po súčasnosť.

## **Kľúčové slová**

Google, vyhľadávače, PageRank, ranking, technológia, hodnotenie stránok, vyhľadávanie na internete

## **Abstract**

The thesis describes a search engine of Google company and its method of web pages evaluating, which is used for sorting them on the search engine result page. In addition to the description of how search engines work in general, the thesis targets on differences between Google and other search engines, as well as it mentions characteristics and principles designed by its founders, Larry Page and Sergey Brin, which make it unique. In the second part there are outlined some of many signals which are rated for every single web page. The most concerned signal is PageRank as the leading idea of pages evaluating, based on their link scheme which is founded on principles of citation analysis. Factors for evaluating are analysed using particular examples, and the topic is ended by the description of Google search engines changes from its beginning in 1998 until present.

## **Keywords**

Google, search engines, pagerank, ranking, technology, webpages evaluating, web searching

# OBSAH

1	ÚVOD .....	12
2	SPOLOČNOSŤ GOOGLE .....	13
2.1	O spoločnosti .....	13
2.2	Založenie spoločnosti .....	13
3	WEBOVÉ VYHLÁDÁVAČE.....	15
3.1	Úvod do vyhľadávania (1994 – 2000) .....	15
3.2	Štruktúra vyhľadávačov .....	15
3.3	Vyhľadávací proces: prezeranie, indexovanie a tvorba požiadavky vyhľadávania..	16
4	VYHLÁDÁVAČ GOOGLE.....	19
4.1	Dôvod vzniku Google .....	19
4.2	Čím je Google významný oproti iným SE.....	20
4.3	Vlastnosti systému .....	21
4.3.1	<i>PageRank</i> .....	21
4.3.2	<i>Anchor text</i> .....	29
4.3.3	<i>Ďalšie vlastnosti</i> .....	30
4.4	Faktory ovplyvňujúce poradie výsledkov na SERP .....	30
4.4.1	<i>Faktory týkajúce sa domény</i> .....	30
4.4.2	<i>Faktory na úrovni stránky</i> .....	31
4.4.3	<i>Faktory na úrovni webového sídla</i> .....	33
4.4.4	<i>Faktory súvisiace so spätnými odkazmi</i> .....	34
4.4.5	<i>Faktory interakcie používateľov</i> .....	36
4.4.6	<i>Špeciálne pravidlá algoritmu</i> .....	36
4.4.7	<i>Sociálne signály</i> .....	37
4.4.8	<i>Signály týkajúce sa značky</i> .....	38
4.4.9	<i>Spam na stránke</i> .....	38
4.4.10	<i>Spam v okolí stránky</i> .....	39
5	UKÁŽKA HODNOTIACICH SIGNÁLOV NA PRAKTICKOM PRÍKLADE .....	40
5.1	Úvod do výskumu .....	40
5.2	Výber kľúčového slova .....	40
5.3	Výber odkazov .....	42
5.4	Výber hodnotiacich faktorov .....	44

5.5	Analýza .....	44
5.5.1	<i>Analýza faktorov na úrovni domény</i> .....	44
5.5.2	<i>Analýza faktorov na úrovni stránky</i> .....	45
5.5.3	<i>Analýza faktorov na úrovni webového sídla</i> .....	46
5.5.4	<i>Analýza faktorov súvisiacich so spätnými odkazmi</i> .....	47
5.5.5	<i>Analýza špeciálnych pravidiel algoritmu</i> .....	48
5.5.6	<i>Analýza sociálnych signálov</i> .....	48
5.5.7	<i>Analýza signálov týkajúcich sa značky</i> .....	48
5.5.8	<i>Vyhodnotenie</i> .....	49
6	NAJVÝZNAMNEJŠIE ZMENY ALGORITMU .....	51
6.1	Aktualizácie v rokoch 1998 - 2002 .....	51
6.2	Aktualizácie roku 2003 .....	52
6.2.1	<i>Boston</i> .....	52
6.2.2	<i>Cassandra</i> .....	52
6.2.3	<i>Dominic</i> .....	53
6.2.4	<i>Esmeralda</i> .....	53
6.2.5	<i>Fritz</i> .....	54
6.2.6	<i>Supplemental index</i> .....	54
6.2.7	<i>Florida</i> .....	54
6.3	Aktualizácie roku 2004 .....	55
6.3.1	<i>Austin</i> .....	55
6.3.2	<i>Brandy (známa aj ako Vince)</i> .....	55
6.4	Aktualizácie roku 2005 .....	56
6.4.1	<i>NoFollow</i> .....	56
6.4.2	<i>Allegra</i> .....	56
6.4.3	<i>Bourbon</i> .....	57
6.4.4	<i>Sitemaps XML</i> .....	57
6.4.5	<i>Personalized Search</i> .....	58
6.4.6	<i>Gilligan</i> .....	58
6.4.7	<i>Google Local / Maps</i> .....	58
6.4.8	<i>Jagger</i> .....	58
6.4.9	<i>Big Daddy</i> .....	59
6.5	Aktualizácie roku 2006 .....	59
6.5.1	<i>Supplemental</i> .....	59

6.5.2	<i>False Alarm</i> .....	59
6.6	Aktualizácie roku 2007 .....	60
6.6.1	<i>Googlebomb</i> .....	60
6.6.2	<i>Universal Search (aka Google 2.0)</i> .....	60
6.6.3	<i>Buffy</i> .....	60
6.7	Aktualizácie roku 2008 .....	60
6.7.1	<i>Dewey</i> .....	60
6.7.2	<i>Google Suggest</i> .....	61
6.8	Aktualizácie roku 2009 .....	61
6.8.1	<i>Vince</i> .....	61
6.8.2	<i>Rel-Canonical Tag</i> .....	61
6.8.3	<i>Caffeine</i> .....	61
6.8.4	<i>Real-Time Search</i> .....	62
6.9	Aktualizácie roku 2010 .....	62
6.9.1	<i>Google Places</i> .....	62
6.9.2	<i>May Day</i> .....	62
6.9.3	<i>Caffeine (Rollout)</i> .....	63
6.9.4	<i>Brand update</i> .....	63
6.9.5	<i>Google Instant</i> .....	63
6.9.6	<i>Instant Previews</i> .....	64
6.9.7	<i>Social Signals</i> .....	64
6.9.8	<i>Negative Reviews</i> .....	64
6.10	Aktualizácie roku 2011 .....	64
6.10.1	<i>Penalizácia Overstock.com</i> .....	64
6.10.2	<i>Attribution Update</i> .....	65
6.10.3	<i>Update Panda (známa aj ako aktualizácia Farmer)</i> .....	65
6.10.4	<i>Tlačidlo +1</i> .....	65
6.10.5	<i>Panda 2.0</i> .....	65
6.10.6	<i>Panda 2.1</i> .....	65
6.10.7	<i>Schema.org</i> .....	66
6.10.8	<i>Panda 2.2</i> .....	66
6.10.9	<i>Google+</i> .....	66
6.10.10	<i>Panda 2.3</i> .....	66
6.10.11	<i>Panda 2.4</i> .....	66

6.10.12	<i>Rozšírenie vnorených odkazov (Sitelinks)</i> .....	68
6.10.13	<i>Prvky číslovania stránok</i> .....	68
6.10.14	<i>516 aktualizácií</i> .....	68
6.10.15	<i>Panda 2.5</i> .....	68
6.10.16	<i>Panda „Flux“</i> .....	68
6.10.17	<i>Šifrovanie požiadaviek</i> .....	68
6.10.18	<i>Aktualizácia čerstvosti (Freshness update)</i> .....	69
6.10.19	<i>Balíček 10 aktualizácií</i> .....	69
6.10.20	<i>Panda 3.1</i> .....	69
6.10.21	<i>Decembrový balíček 10 aktualizácií</i> .....	69
6.11	<i>Aktualizácie roku 2012</i> .....	69
6.11.1	<i>Januárový balíček 30 aktualizácií</i> .....	69
6.11.2	<i>Search + Your World</i> .....	69
6.11.3	<i>Panda 3.2</i> .....	70
6.11.4	<i>Reklamné odkazy nad prehybom stránky (Ads above the Fold)</i> .....	70
6.11.5	<i>Venice</i> .....	70
6.11.6	<i>Balíček 40 aktualizácií</i> .....	70
6.11.7	<i>Panda 3.3</i> .....	71
6.11.8	<i>Video o kvalite vyhľadávania</i> .....	71
6.11.9	<i>Panda 3.4</i> .....	71
6.11.10	<i>Trest za prehnané optimalizovanie stránky</i> .....	71
6.11.11	<i>Marcový balíček 50 aktualizácií</i> .....	71
6.11.12	<i>Chyba v hodnotení parkovaných domén (Parked Domain Bug)</i> .....	71
6.11.13	<i>Panda 3.5</i> .....	72
6.11.14	<i>Penguin</i> .....	72
6.11.15	<i>Panda 3.6</i> .....	72
6.11.16	<i>Aprílový balíček aktualizácií</i> .....	72
6.11.17	<i>Knowledge Graph</i> .....	74
6.11.18	<i>Penguin 1.1</i> .....	74
6.11.19	<i>Májový balíček aktualizácií</i> .....	74
6.11.20	<i>Panda 3.7</i> .....	74
6.11.21	<i>Panda 3.8</i> .....	74
6.11.22	<i>Upozornenia na odkazy (Links Warnings)</i> .....	75
6.11.23	<i>Panda 3.9</i> .....	75



6.11.24	<i>Balíček 86 aktualizácií za obdobie jún/júl .....</i>	75
6.11.25	<i>Penalizácia DMCA.....</i>	75
6.11.26	<i>Stránka vyhľadávača so 7 výsledkami .....</i>	75
6.11.27	<i>Panda 3.9.1 .....</i>	76
6.11.28	<i>Panda 3.9.2 .....</i>	76
6.11.29	<i>Panda #20 .....</i>	76
6.11.30	<i>Aktualizácia domén s presnou zhodou (EMD).....</i>	76
6.11.31	<i>Balíček updatov august/september.....</i>	76
6.11.32	<i>Penguin 3 .....</i>	76
6.11.33	<i>Rozloženie stránky #2.....</i>	77
6.11.34	<i>Panda #21 .....</i>	77
6.11.35	<i>Panda #22 .....</i>	77
6.11.36	<i>Rozšírenie Knowledge Graph .....</i>	77
6.11.37	<i>Panda #23 .....</i>	77
6.12	<i>Aktualizácie roku 2013 .....</i>	77
6.12.1	<i>Panda #24.....</i>	77
6.12.2	<i>Panda #25.....</i>	78
6.13	<i>Budúcnosť webového vyhľadávania .....</i>	78
7	<i>ZÁVER .....</i>	83
8	<i>ZOZNAM OBRÁZKOV V TEXTE.....</i>	84
9	<i>ZOZNAM POUŽITÝCH ZDROJOV .....</i>	85
10	<i>ZOZNAM PRÍLOH .....</i>	90
11	<i>PRÍLOHY</i>	

## **Zoznam použitých skratiek**

<b>CTR</b>	Click Through Rate
<b>DMCA</b>	Digital Millennium Copyright Act
<b>DMOZ</b>	<a href="http://directory.mozilla.org">directory.mozilla.org</a>
<b>EMD</b>	Exact Match Domain
<b>HTML</b>	HyperText Markup Language
<b>IP</b>	Internet Protocol
<b>KW</b>	Keyword
<b>LBC</b>	Local Business Center
<b>LSI</b>	Latent Semantic Indexing
<b>OOP</b>	Over Optimization Penalty
<b>PR</b>	PageRank
<b>RSS</b>	Rich Site Summary
<b>SEO</b>	Search Engine Optimization
<b>SERP</b>	Search Engine Result Page
<b>TBPR</b>	Toolbar PageRank
<b>TLD</b>	Top Level Domain
<b>URL</b>	Uniform Resource Locator
<b>WWW</b>	World Wide Web

## **PREDHOVOR**

Téma tejto diplomovej práce bola zvolená na základe potreby zdrojov tejto problematiky, ktoré sú v súčasnosti dostupné prevažne na zahraničných stránkach. Tematikou sa zaoberá obrovské množstvo zdrojov, mojím cieľom bolo tieto zdroje preskúmať a na základe toho vytvoriť prehľad celej problematiky – od úplných základov a princípov fungovania webového vyhľadávania, až po konkrétne aktualizácie v rámci vyhľadávača Google.

Úvod práce je teda zoznámením so spoločnosťou Google a jej zakladateľmi, v nasledujúcej kapitole som už prešla priamo k vyhľadávaciemu procesu a jeho princípom a popisu činností, ktoré musia byť vykonané na to, aby mohli byť na vyhľadávaciu požiadavku zadanú do okna vyhľadávača zobrazené čo najrelevantnejšie výsledky v príslušnom poradí.

K ťažnej téme celej práce, a to k vyhľadávaniu v Google, som prešla cez popis dôvodov jeho vzniku a odlišností, ktorými vyniká spomedzi iných nástrojov podobného typu. Okrem PageRanku, ako hodnotiaceho faktoru, ktorému je venovaná samostatná podkapitola, som sa zaoberala aj niektorými ďalšími z obrovskej množiny faktorov vplyvajúce na hodnotenie stránok a ich následne zoradenie na stránke s výsledkami hľadania.

Od vymenovania týchto faktorov rozdelených do logických skupín som prešla k praktickej časti, v ktorej som na príklade zobrazených výsledkov pri vyhľadávaní vopred zvoleného kľúčového slova ukázala jednotlivé faktory. Pretože som bola limitovaná dostupnosťou dát, hodnotila som iba výber približne 60 – 65 faktorov. Tento výskum ukázal, že algoritmus Google pre hodnotenie stránok je zložitým procesom, ktorý sa neustále mení, faktorov, ktoré hodnotí je obrovské množstvo a každý z nich má inú veľkosť vplyvu na hodnotenie.

Posledná, aj keď nie menej dôležitá kapitola obsahuje popis najvýznamnejších zmien a aktualizácií algoritmu a princípov hodnotenia a usporadúvania stránok, od začiatku vyhľadávača Google v roku 1998 až po najnovšie updaty tohto roku.

# ÚVOD

Počiatky webových vyhľadávačov siahajú už do doby vzniku webu, ktorého veľkosť sa vyvíjala fenomenálnou rýchlosťou. Koncom roku 1994 obsahoval index jedného z prvých vyhľadávačov až 110 000 webových stránok.

Princíp vyhľadávania ale nespočíva len v zadaní vyhľadávacej požiadavky do okna vyhľadávača, pred tým je potrebné, aby roboty prešli obrovské množstvo webov a zaradili ich stránku do svojho indexu, kde sa im pridávajú identifikačné údaje spolu s hodnotením ich dôležitosti. Práve v tejto časti procesu získavajú aj známy PageRank, číselnú hodnotu zo stupnice od 1 do 10, ktorá im je pridelovaná predovšetkým na základe ich počtu odkazov. Tento princíp je založený na citačnej analýze, pri ktorej bola na základe citácií, ktorými boli dokumenty navzájom poprepájané, vytvorená sieť informácií, ktorej boli vyjadrené vzťahy medzi týmito dokumentmi. Citovaním určitého diela dával autor tomuto dielu hlas, čo bolo znakom kvality diela. V prípade webových stránok je hlasom odkaz, ktorým stránka inej stránke zvyšuje PageRank a stránka tak naberá na vážnosti.

Hoci sa PageRank stal prielomovým spôsobom hodnotenia stránok, nie je to jediný faktor, ktorý Google zohľadňuje pri hodnotení stránok. V práci je uvedených viac ako dvesto faktorov, ktoré majú na toto hodnotenie väčší alebo menší vplyv, ich vážnosť sa v čase mení a s vývojom internetového prostredia a príchodom stále ďalších možností, sa tento zoznam ďalej rozrastá.

Cieľom spoločnosti Google je neustále zlepšovanie technológie, na základe ktorej by boli na vyhľadávacie požiadavky zobrazené čo najrelevantnejšie, najpodstatnejšie a najčerstvejšie informácie a dokumenty, a preto na vývoji svojho algoritmu bez prestávky pracuje.

Detailný popis spôsobu svojho hodnotenia si ale chráni. Na základe dostupných informácií sme síce schopný pochopiť myšlienku tohto hodnotenia, dokonca je známa aj pôvodná formula výpočtu PageRanku a jej mnohé variácie, presný proces však pozná len sám Google.

Pri spracovávaní výsledkov svojho bádania a výskumu som narazila na problém nejednotnosti terminológie tejto oblasti v slovenskom jazyku, v práci som teda používala výrazy, ktoré boli v danom kontexte najvhodnejšie a prílohy obsahujú slovníček vysvetlených pojmov spolu s anglickými aj slovenskými ekvivalentmi.

# 1 SPOLOČNOSŤ GOOGLE

## 1.1 *O spoločnosti*

Spoločnosť Google sa primárne zameriava na rýchle, jednoduché a kvalitné sprostredkovanie najrelevantnejších dostupných informácií čo najväčšiemu počtu užívateľov. Technologické inovácie priniesli spoločnosti niekoľko ocenení vrátane ceny WIRED Readers Raves v kategórii Most Intelligent Agent (najinteligentnejší sprostredkovateľ, obchodník, tvorca), ceny 2000 Webby and People's Voice Award For Best Technical Achievement (cena Hlasu siete a ľudí za najlepšie technický úspech), ceny Best Search Engine on the Internet (najlepší vyhľadávač na internete) od časopisu Yahoo! Internet Life, ceny Top Ten Best Cybertech roku 1999 od časopisu Time (10 najlepších kybernetických technológií, Technical Excellence Award (ocenenie vynikajúcej technickej kvality) od PC Magazine a ceny Best Search Engine (najlepší vyhľadávač) od The Net. Rastúci počet spoločností, vrátane Yahoo!, AOL (netscape), Cisco, Washingtonpost.com a Red Hat, sa spoliehalo práve na Google, aby posilnil vyhľadávanie na ich webových stránkach. (Google Launches The Google Toolbar 2012)

## 1.2 *Založenie spoločnosti*

Zakladatelia spoločnosti Google, Sergey Brin a Larry Page, sa stretli v roku 1995, keď Page navštívil cez náborový víkend oddelenie počítačovej vedy Standfordskej univerzity. Brin, ktorý bol v tom čase absolventom druhého ročníka, pomáhal ako sprievodca nových záujemcov a Page sa dostal do jeho skupiny. Počas svojho prvého stretnutia prebrali mnoho tém a takmer na všetkom sa zhodli. Krátko na to, ako začal svoje štúdium na Standforde, začal Page pracovať na webovom projekte, spočiatku nazývanom BackRub, ktorý využíval štruktúru odkazov webu. Brinovi pripadala Pageova práca zaujímavá, a tak začali spoločne pracovať na projekte, ktorý neskôr absolútne zmenil vyhľadávanie na webe. (Wills 2006)

Brin a Page si uvedomovali, že pracujú na vyhľadávači, ktorý sa prispôsobí neustále sa zväčšujúcemu webu, a tak zmenili jeho názov z BackRub na Google (prešmyčka bežne používaného výrazu „googol“, čo je číslo 10<sup>100</sup>). Neboli však schopní presvedčiť existujúce spoločnosti webových vyhľadávačov, aby používali technológiu, ktorú vytvorili, aj napriek

tomu, že bola lepšia ako ich doterajšie, a tak sa Brin a Page rozhodli založiť svoju vlastnú spoločnosť. S finančnou pomocou malej skupiny počiatočných investorov založili v septembri 1998 spoločnosť a webový vyhľadávač Google, Inc.

Všeobecná verejnosť takmer okamžite zaznamenala to, čo Brin, Page a ďalší z akademickej komunity webového vyhľadávania už vedeli, a síce, že vyhľadávač Google prináša oveľa kvalitnejšie výsledky ako ostatné vyhľadávače. Tie sa pri určovaní poradia vo výsledkoch vyhľadávania spoliehali výhradne na obsah webových stránok a Brin a Page si uvedomovali, že vývojári webových stránok môžu ľahko ovplyvňovať výsledky vyhľadávania umiestňovaním skrytých informácií na svoje webové stránky. Vytvorili teda algoritmus pre ranking nazvaný PageRank po Larrym Pageovi, ktorý využíval štruktúru odkazov webu pre určovanie dôležitosti webových stránok. Počas spracovávania požiadavky, vyhľadávací algoritmus Googlu kombinoval vopred vypočítané skóre PageRanku s hodnotami textovej zhody pre získanie celkového počtu bodov pre každú webovú stránku. (Wills 2006)

## 2 WEBOVÉ VYHLÁDÁVAČE

### 2.1 Úvod do vyhľadávania (1994 – 2000)

Technológia webových vyhľadávačov zaznamenala obrovský pokrok, aby udržala krok s rýchle rastúcim webom. V roku 1994 mal jeden z prvých webových vyhľadávačov, World Wide Web, index obsahujúci 110 000 webových stránok a iných dokumentov prístupných cez web. V novembri 1997 oznámili popredné webové vyhľadávače, že je zaindexovaných 2 až 100 miliónov webových dokumentov. Predvídalo sa, že do roku 2000 bude postupne doplňovaný súhrnný index obsahovať viac než bilión dokumentov. V tom istom čase neuveriteľne stúpol aj počet vyhľadávaných hesiel vo vyhľadávačoch. V marci a apríli 1994 dosiahol www priemerne 1 500 vyhľadávaní denne. V novembri 1997 Altavista vyhlásil, že vybavuje približne 20 miliónov vyhľadávacích požiadaviek denne. S narastajúcim počtom webových používateľov a automatizovanými systémami vyhľadávačov pre spracovávanie používateľských požiadaviek sa predpokladal do roku 2000 nárast až na stovky miliónov požiadaviek denne. (Brin 2008)

### 2.2 Štruktúra vyhľadávačov

Vyhľadávače sa skladajú zo štyroch hlavných častí:

- *Robot* (spider, crawler) je program používaný na prezeranie internetu, aby boli nájdené a identifikované nové alebo zmenené webové sídla. Následne sú informácie o nich poskytnuté indexačnému mechanizmu vyhľadávača. Populárnejšie weby sú niekedy prezerané detailnejšie a častejšie ako malé weby. Na základe tohto procesu sú webové sídla identifikované svojou URL adresou a ak nie sú zaradené medzi spam, poprípade nevhodné z iného dôvodu, budú zaindexované a pridané do databázy.
- *Indexačný program a index* – hneď ako robot identifikuje stránku, indexačný program zaindexuje obsah stránky, URL, metatagy, URL adresy odkazov na danej stránke a súborové názvy obrázkov.
- *Samotný vyhľadávač* je program, ktorý tieto stránky získava z databázy na základe kritérií požiadavky užívateľa pri vyhľadávaní. V tomto procese je veľmi dôležitým

aspektom poradie, v akom sú výsledky hľadania zobrazené. Algoritmus určovania relevancie (relevance-ranking) závisí od mnohých faktorov, ako je popularita stránky (podľa počtu na ňu odkazujúcich stránok), počet výskytov hľadaného výrazu na stránke, ich blízkosť, umiestnenie (napr. keď sa výraz nachádza v názve, má vyššie hodnotenie, ako výraz v texte a pod.) a mnoho ďalších.

- *HTML rozhranie*, ktoré preberá dotaz a kritériá vyhľadávania od užívateľa – domovská stránka (HomePage) vyhľadávacej služby a stránky pokročilého vyhľadávania, teda všetky časti, ktoré si vieme predstaviť pre konkrétny vyhľadávač – okná vyhľadávania, linky na rôzne dostupné databázy (hľadať v obrázkoch, správach, atď.) a ďalšie funkcie. (Hock c2007, s. 63)

### ***2.3 Vyhľadávací proces: prezeranie, indexovanie a tvorba požiadavky vyhľadávania***

Vyhľadávací proces nezačína zadáním hľadaného hesla do okna vyhľadávača, pred tým je nevyhnutné, aby bolo vykonané prezeranie a indexovanie obrovskej siete dokumentov, čo tvorí dôležitý základ celého vyhľadávania. Podstatou týchto činností je tvorba indexu, vďaka ktorému je vyhľadávač schopný na základe pozbieraných a usporiadaných informácií poskytnúť relevantnú odpoveď na vyhľadávaciu požiadavku. (How Search Works 2012)

Prezeranie voľne dostupných webových stránok v Google vykonáva niekoľko distribuovaných robotov (nazývaných aj „spiders“. Najznámejší z nich je nazývaný Googlebot).

Na tento proces používa spoločnosť Google obrovské množstvo počítačov, ktoré prehľadávajú miliardy internetových stránok. Googlebot používa algoritmický proces: počítačový program určí, ktoré stránky majú byť prechádzané, ako často, a tiež počet, koľko stránok má byť preskúmaných z každého webového sídla. (Googlebot 2012)

URL server posiela zoznam URL adries, ktoré majú byť k dispozícii robotom.

Proces prechádzania stránok Googlebotom začína zoznamom URL adries webových stránok, ktorý je generovaný na základe predchádzajúceho prechádzania stránok a jeho následného rozšírenia pomocou dát z máp webov, ktoré sú poskytované webmastermi. Keď Googlebot navštívi každé z týchto webových sídel, zistí adresu na každú stránku (SRC a HREF) a pridá



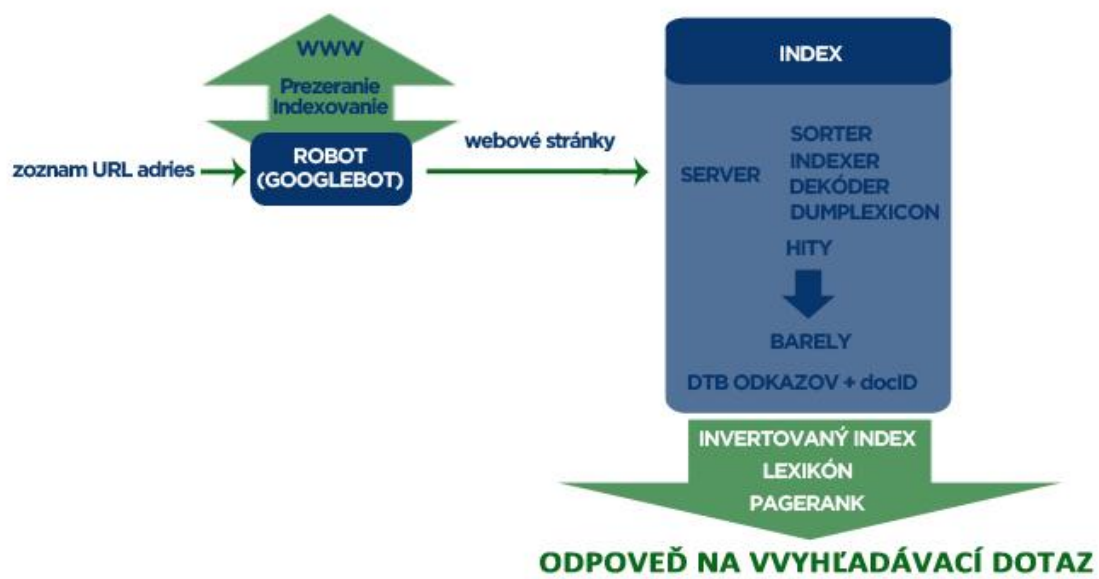
ju do svojho zoznamu stránok na prechádzanie. Nové webové lokality, zmeny v existujúcich a mŕtve odkazy sú zaznamenávané a použité pri aktualizácii Google indexu. (Googlebot 2012)

Webové stránky, ktoré roboty prevezmú, sú následne zaslané na úložný server. Ten potom skomprimuje a uloží webové stránky do repozitára. Každá stránka je označená identifikačným číslom, nazývaným docID, ktoré jej je priradené zakaždým, ako je nová URL adresa z webovej stránky získaná. Indexačnú funkciu má na starosti indexer a sorter. Indexer plní niekoľko funkcií – číta obsah repozitára, dekomprimuje a analyzuje dokumenty. Každý dokument je potom konvertovaný na skupinu výskytov slov nazývaných hity.

Hity zaznamenávajú slovo, jeho pozíciu v dokumente, približnú veľkosť fontu a používanie kapitálok. Indexer distribuuje tieto hity do setov nazývaných barely a vytvára čiastočne usporiadaný forward index. Indexer však vykonáva aj ďalšiu dôležitú funkciu, ktorou je analýza všetkých odkazov danej webovej stránky a ukladá o nich dôležité informácie v anchor súbore. Tento súbor obsahuje dostatok informácií na určenie, kde ktorý odkaz odchádza a prichádza, spolu s textom odkazu.

URL dekodér potom prečíta anchor súbor, skonvertuje relatívne URL adresy na absolútne a premení ich na docID. Anchor text vloží do forward indexu, spolu s docID, na ktoré sa viaže. Zároveň vygeneruje databázu odkazov, ktoré sú spárované s docID. Odkazy v databáze sa potom použijú na výpočet PR všetkých dokumentov.

Sorter vezme barely, ktoré sú usporiadané podľa docID a ukladá ich podľa wordID, aby vygeneroval invertovaný index. Tento proces je vykonávaný na jednom mieste a preto celá operácia potrebuje dočasný priestor. Sorter zároveň vytvára zoznam príslušných WordID pre invertovaný index. Program s názvom DumpLexicon tento zoznam vezme spolu s lexikónom, ktorý vytvoril indexer, a vygeneruje nový lexikón pre použitie vyhľadávača. Ten je spustený webovým serverom a používa lexikón, ktorý vytvoril DumpLexicon spolu s invertovaným indexom a PageRankmi pre zodpovedanie vyhľadávacieho dotazu. (Brin 2008)



Obr. č. 1: Proces tvorby odpovede na užívateľskú požiadavku

## 3 VYHLÁDÁVAČ GOOGLE

### 3.1 *Dôvod vzniku Google*

Začiatkom 90. rokov so vznikom prvých nástrojov na vyhľadávanie na webe panoval názor, že „najlepšia navigačná služba je tá, ktorá na Webe nájde takmer všetko“. Roku 1997 sa však ukázalo, že úplnosť indexu nie je jediný faktor, ktorý by zaručil kvalitu vyhľadaných výsledkov. Množstvo relevantných výsledkov nebolo zobrazených na úkor nekvalitných výsledkov, a zo štyroch, v tom čase najvýznamnejších vyhľadávačov, sa len jeden objavoval vo svojich vlastných výsledkoch po zadaní jeho názvu do vyhľadávacieho okna.

Jednu z hlavných príčin problémov videli Brin a Page v neustálom raste počtu dokumentov, pričom schopnosť užívateľov na ich prehliadanie zostávala nezmenená a zaujímalo ich predovšetkým prvých 10 nájdených odkazov. Tým vznikla potreba nástroja, ktorý by bol schopný medzi tieto odkazy začleniť len relevantné výsledky<sup>1</sup> pre danú požiadavku. Vysoká presnosť je nevyhnutná hoci aj na úkor počtu nájdených dokumentov.

Zakladatelia spoločnosti Google v tej dobe verili, že rozšírenie použitia hypertextových informácií môže vyhľadávaniu pomôcť, predovšetkým štruktúra a texty odkazov poskytovali dostatok informácií pre určenie ich relevancie a kvality, čo neskôr využili v hodnotení Google.

Popri rastu obsahu sa Web stával čoraz komerčnejším. Počet serverov na doménach .com sa od roku 1993 zvýšil z 1,5% na neuveriteľných 60% do roku 1997 a migrovali na ňu aj vyhľadávače z akademických domén. Väčšina vyhľadávačov v tom čase patrila firmám a boli prevažne reklamne orientované. Cieľom tvorcov Google bolo posunutie vývoja viac do akademickej sféry.

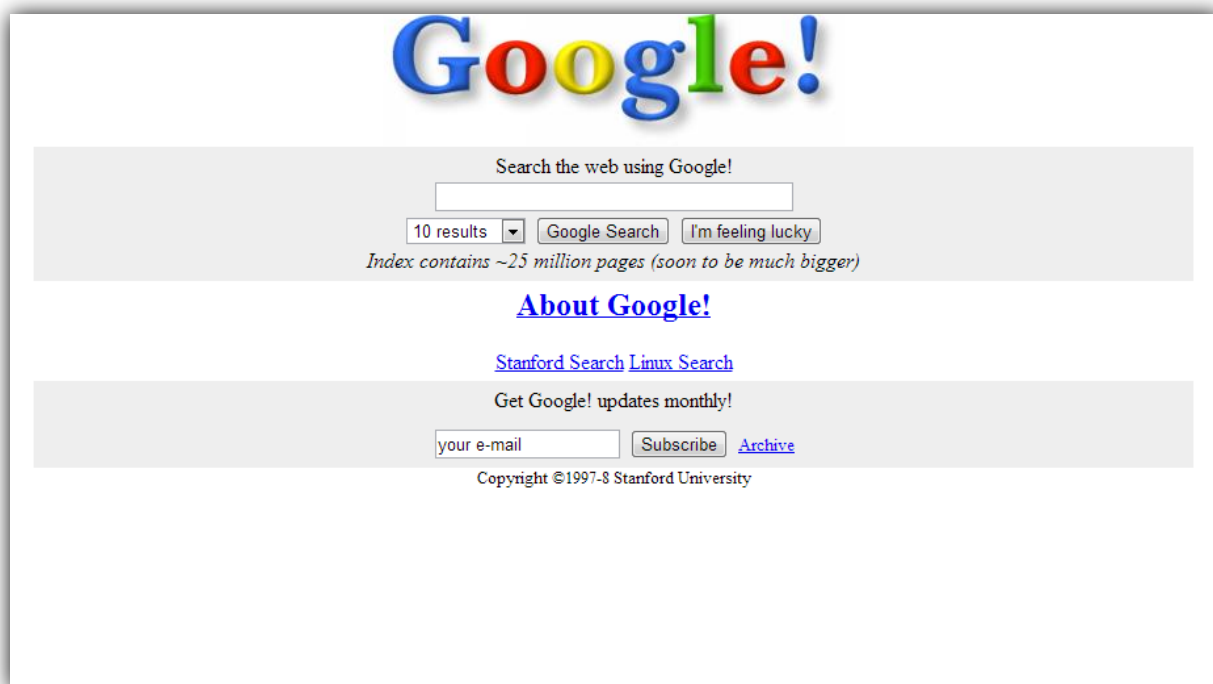
Ďalším argumentom pre vytvorenie nového vyhľadávacieho systému bola jeho použiteľnosť veľkým počtom užívateľov a na ňu zameraný výskum – denne boli totiž vyhľadávané desiatky miliónov hesiel, o ktorých však neboli dostupné informácie kvôli ich komerčnej hodnote.

Page a Brin plánovali vytvorenie architektúry podporujúcej dovtedy neznámy výskum obrovskej množiny webových dát. Google zhromažďoval všetky dokumenty, ktoré prezrel,

---

<sup>1</sup> Za relevantné výsledky Brin a Page považovali podľa ich slov „najlepšie dokumenty, spomedzi desiatok tisícov menej relevantných dokumentov“.

v komprimovanej forme, aby vytvoril prostredie, do ktorého by výskumníci mohli ihneď vstúpiť, spracovával obrovské časti webu a prinášal zaujímavé výsledky, ktoré by iným spôsobom nebolo možné získať. Okrem toho chceli vytvoriť prostredie, v ktorom by vedci alebo študenti mohli na týchto rozsiahlych internetových dátach robiť zaujímavé pokusy. (Brin 2008)



Obr.č. 2: Domovská stránka vyhľadávača Google po jeho spustení (11.11.1998)

(Zdroj: <http://web.archive.org/web/19981111183552/http://google.stanford.edu/>)

### 3.2 Čím je Google významný oproti iným SE

Pred Googlom internetové vyhľadávače používali pre hodnotenie a usporiadanie webových stránok vo výsledkoch vyhľadávania to, nakoľko vyhľadaná stránka odpovedala používateľskému dotazu. (Houser 2008)

Google sa vyznačoval dvomi vlastnosťami, vďaka ktorým bol schopný priniesť veľmi presné výsledky: používal odkazovú štruktúru webu pre výpočet hodnoty kvality - PageRanku pre každú webovú stránku a využíval odkazy na zlepšenie vyhľadaných výsledkov. (Brin 2008)

Ako odpoveď na požiadavku používateľa dáva na prvé miesta stránky, ktoré sú najdôležitejšie a najvýznamnejšie; samozrejme berie do úvahy obsah stránky vzhľadom na slová, ktoré používateľ hľadá. Hodnota PR je teda významná najmä u požiadaviek všeobecného typu.

Okrem PR sa Google od konkurencie odlišil aj inými vecami, napríklad neportálovým prístupom, preferenciou textovej reklamy alebo platbou za kliknutie. (Houser 2008)

### **3.3 Vlastnosti systému**

#### **3.3.1 PageRank**

##### **3.3.1.1 Google Toolbar PageRank**

Google Toolbar PageRank a PageRank sú dve rôzne nesynchronizované metriky založené na počte a kvalite odkazov, ktoré tvoria celkové hodnotenie danej stránky. (Google PageRank Update History 2012)

Hodnota PR webovej stránky korešponduje s údajom vektora PR  $\pi$ . Pretože je  $\pi$  vektorom pravdepodobnosti distribúcie, všetky prvky  $\pi$  sú nezáporné a ich súčet dáva 1. Panel nástrojov Google obsahuje funkciu zobrazenia PR, ktorá poskytuje údaj (odhad) PR práve navštívenej webovej stránky. Hodnoty PR na paneli nástrojov predstavujú celé čísla od 0 (najnižší) do 10 (najvyšší). Hoci niektorí experti na optimalizáciu pre vyhľadávače podceňujú presnosť hodnôt na paneli nástrojov, webová stránka Googlu v súvislosti s funkciou panela nástrojov uvádza:

„Zobrazenie PR: zaujíma Vás, či nový web stojí za Váš čas? Použite PR zobrazený na paneli nástrojov, aby ste zistili, ako Google algoritmus hodnotí dôležitosť stránky, ktorú si prezeráte.“

Výsledky Google panela nástrojov, ktoré priniesol Google pre vyhľadávanie, ukazujú, že veľa ľudí venuje hodnotám PR panela nástrojov značnú pozornosť. Jedna webová stránka uvádza, že vlastníci webových stránok sa na PR paneli nástrojov stali závislými.

Hoci Google nevysvetľuje, podľa čoho sú hodnoty PR panela nástrojov určované, je možné že sú založené na logaritmickej stupnici. Je ľahké overiť, že niekoľko webových stránok získalo skóre PR panela nástrojov 10, no mnoho webových stránok má len veľmi nízke skóre.

PR je síce dôležitou súčasťou celkového hodnotenia výsledkov v Google, ale nie je to jeho jediná zložka. Pretože vysoké skóre PR webovej stránky negarantuje, že sa v zozname výsledkov umiestni vysoko, experti na optimalizáciu pre vyhľadávače zdôrazňujú, že onpage faktory, ako je umiestnenie a frekvencia dôležitých slov, musia byť zohľadnené pri tvorbe kvalitných webových stránok. Dokonca aj spravodajské médiá sa začali zameriavať na úpravu nadpisov a textov svojich článkov kvôli zlepšeniu hodnotenia vo výsledkoch vyhľadávača. Faktom je, že väčšina používateľov webových vyhľadávačov očakáva, že rýchlo nájde relevantné výsledky pre akúkoľvek tému. Pre uspokojenie tejto požiadavky si musí byť Google istý, že sa medzi prvými výsledkami vyhľadávania objavia najrelevantnejšie webové stránky. V rámci zachovania konkurenčného prostredia preto musia spoločnosti a spravodajské médiá zisťovať spôsoby, ako sa tam dostať. (Wills 2006)

### **3.3.1.2 Definícia**

Zjednodušene povedané, PageRank je číselná hodnota, ktorá predstavuje dôležitosť stránky na webe. Google ho vypočítava tak, že ak jedna stránka odkazuje na inú, dáva jej tým svoj hlas. Čím viac hlasov stránka dostane, tým je stránka dôležitejšia. Na druhej strane, odkazujúca stránka určuje, akú dôležitosť bude mať hlas, ktorý danej stránke prepožičiava. Google túto dôležitosť počíta z hlasov, ktoré získala odkazujúca stránka. Dôležitosť jednotlivých hlasov teda určuje vypočítaný PR každej stránky. (Craven 2003)

PR predstavuje exponenciálnu číselnú stupnicu od 0 do 10, čo znamená, že čím viac odkazov stránka má, tým vyššiu hodnotu má jej PR. Každý odkaz však PR ovplyvňuje inou mierou (napríklad odkazy z registrov buď započítané nie sú, alebo predstavujú nižšiu hodnotu ako odkazy z iných webových stránok) a neexistuje jasný a jednoznačný vzorec na jeho vypočítanie.

Približná hodnota PR sa však dá odhadnúť – PR1 pri 10 odkazoch, PR2 11-100 odkazov, PR3 101-1 000 odkazov, PR4 1001-10 000 odkazov a podobne (napríklad Wikipedia.org a google.com majú PR 9, PR 10 sa vyskytuje len v zriedka). Zjednodušene sa dá povedať, že čím vyšší má stránka PR, tým vyššie je umiestnená vo výsledkoch vyhľadávania, a čím vyšší PR má odkaz smerujúci na určitú stránku, tým väčšiu vážnosť jej pridá.

Ak má teda stránka PR 2 a viac, je to dost na to, aby sa zobrazovala medzi relevantnými výsledkami vyhľadávania v určitej špecifickej oblasti – ak je dobre optimalizovaná pre vyhľadávače (má dobré SEO) a nespadá do skupiny s príliš vysokou konkurenciou. V prípade, že do takejto skupiny patrí, mala by dosiahnuť PR 3, 4 alebo vyšší. Podobne

je to v prípade, ak daná webová stránka ponúka niekoľko rôznych obsahov alebo pokrýva širokú geografickú oblasť, a je pre ňu nevyhnutné získať čo najviac na ňu smerujúcich odkazov namiesto toho, aby sa zobrazovala na popredných pozíciách vo viacerých tematických oblastiach súčasne.

Každá stránka určitého webového sídla môže mať iný PR. Vo všeobecnosti však platí, že domovská stránka bude mať PR o 1 vyšší ako ostatné stránky, na ktoré vedie viac odkazov. Stránky, na ktoré vedú interné odkazy budú mať taktiež nižší PR, pretože na nich neodkazujú iné webové stránky. (List of Google PageRank updates 2011 2011)

PR pomáha Googlu pri rozhodovaní o dôležitosti jednotlivých stránok. Má veľký význam, pretože ovplyvňuje ranking a umiestnenie stránok vo výsledkoch vyhľadávania. Nie je to síce jediný faktor určujúci poradie stránok pri vyhľadávaní, no je jedným z najvýznamnejších. (Craven 2003)

### **3.3.1.3 Význam PageRanku**

Odkazový (resp. citačný) graf webu nebol napriek svojej dôležitosti využívaný v existujúcich internetových vyhľadávačoch, a tak Lawrence Page a Sergey Brin vytvorili mapy zložené až z 518 miliónov odkazov, čo bola významná časť celkového počtu odkazov. Na základe týchto máp bolo možné za krátky čas vypočítať PageRank stránky ako objektívnu metriku jej dôležitosti z pohľadu citovania, čo zodpovedalo subjektívnemu hodnoteniu jej významu používateľmi. Pre väčšinu populárnych tém bolo hľadanie jednoduchého textu s obmedzením na názvy webových stránok vykonávané výborne použitím hodnotenia výsledkov prostredníctvom PageRanku. Jeho význam sa výrazne prejavil aj v fulltextovom vyhľadávaní v hlavnom systéme Google. (Brin 2008)

Hoci celkové poradie vo výsledkoch vyhľadávania určuje niekoľko faktorov, Google pokladá za srdce softvéru svojho vyhľadávača práve PageRank. Zopár rýchlych vyhľadávanií na internete ukáže, že obchodné aj akademické komunity stavajú PageRank veľmi vysoko. Obchodná komunita si je vedomá toho, že Google zostáva prvou voľbou medzi vyhľadávačmi a PageRank zohráva dôležitú úlohu v poradí, v akom sa webové stránky zobrazujú.

Zvyšovanie hodnoty PageRanku sa preto stalo významným komponentom marketingových stratégií spoločností. Akademická komunita spozorovala, že PageRank má spojitosť s niekoľkými oblasťami matematickej a počítačovej vedy, konkrétne teóriou matíc,

numerickými analýzami, vyhľadávaním informácií a teóriou grafov. Výsledkom je mnoho výskumov zameraných na vysvetlenie a vylepšenie PageRanku. (Wills 2006)

Zjednodušene povedané, stránka je tým významnejšia, čím viac stránok na ňu odkazuje, a tiež čím viac významných stránok na ňu odkazuje. Odkazy totiž nemajú rovnakú váhu, tá závisí od toho, na koľko iných stránok daná stránka odkazuje. (Houser 2008)

Algoritmus PageRanku prideluje PR skóre každej z viac než 25 miliárd webových stránok. Modeluje správanie bežného webového používateľa, ktorý si náhodne zvolí webovú stránku, ktorú chce zobraziť, zo zoznamu dostupných stránok. Potom náhodne vyberie odkaz na tejto stránke, ktorý odkazuje na inú webovú stránku. Používateľ pokračuje voľbou odkazov podľa náhodnej postupnosti webových stránok v rámci webového sídla, až kým sa nerozhodne premiestniť na inú webovú stránku iným spôsobom ako je výber odkazu. Rozhodnutie, ktorú webovú stránku má navštíviť ako ďalšiu, nezávisí od už navštívených stránok a vzorový používateľ nikdy neprestane využívať webové stránky. PR stránky teda reprezentuje pravdepodobnosť, že náhodný používateľ webu si na prehliadanie vyberie práve túto stránku. (Wills 2006)

### **3.3.1.4 Výpočet PageRanku**

#### *3.3.1.4.1 Čo Google započítava*

Nie všetky odkazy Google započítava. Napríklad, často odfiltruje odkazy zo známych odkazových fariem a kvôli niektorým odkazom stránky penalizuje. Pripomína, že správcovia webových sídiel síce nemôžu byť zodpovední za odkazy, ktoré na ich stránku odkazujú, no môžu rozhodovať, ktoré odkazy budú z ich stránky odkazovať. Z tohto dôvodu odkazy vedúce na stránku danej stránky uškodiť nemôžu, ale odkazy, ktoré vedú zo stránky, môžu byť pre ňu skutočne škodlivé, v prípade že odkazujú na penalizovaný web, preto treba byť opatrný pri tvorbe odkazov. Ak má určitý web PR 0, pravdepodobne je to kvôli penalizácii a bolo by nerozumné naňho odkázať. (Craven 2003)

#### *3.3.1.4.2 Rovnica výpočtu*

Citovanie literatúry v akademickej oblasti sa premietlo do webového priestoru - výpočet PR určitej stránky bol založený na súčte všetkých jej odkazov (citácií) – či už boli vedené na ňu alebo z nej. (Brin 2008)

Rovnica pre výpočet PR:

$$PR(A) = (1-d) + d (PR(t1) / C(t1) + \dots + PR(tn) / C(tn))$$



Toto je pôvodná rovnica výpočtu PR, ktorá bola zverejnená v čase, keď bol PR vyvíjaný a je pravdepodobné, že Google stále používa jej obdobu.

"t1-tn" – stránky odkazujúce na stránku A

"C" – počet odkazov smerujúcich smerom von z danej stránky

"d" – predstavuje činiteľ tlmenia, ktorý môže nadobúdať hodnoty 0 až 1, väčšinou je nastavený na 0,85 (Craven 2003)

„PR(A)“ – PageRank určitej stránky môže byť vypočítaný použitím jednoduchého iteračného algoritmu a zodpovedá hlavnému vlastnému vektoru normalizovanej odkazovej matice webu.

Hoci táto kalkulácia zahŕňa ešte mnoho ďalších detailov, môže byť PageRank pre 26 miliónov webových stránok vypočítaný na stredne veľkej pracovnej stanici v priebehu niekoľkých hodín. (Brin 2008)

#### 3.3.1.4.3 Názorný popis výpočtu:

PageRank môžeme teda chápať ako model správania náhodného užívateľa, ktorý sa dostane na náhodnú webovú stránku a kliká na ďalšie a ďalšie odkazy bez toho, aby sa vrátil späť, a po určitom čase začne odznova na inej náhodne zvolenej stránke. Pravdepodobnosť, že náhodný užívateľ navštívi určitú stránku, je jej PageRankom. Činiteľ tlmenia d je pravdepodobnosť stránky, že návštevník prestane s klikaním na odkazy na nej a začne odznova na novej náhodne zvolenej stránke. Dôležitý rozdiel je v tom, či je faktor tlmenia priradený len jednej stránke, alebo skupine stránok, čo umožňuje personalizáciu a takmer znemožňuje úmyselné manipulovanie systému za účelom získania vyššieho rankingu.

Určitá stránka môže mať vysoký PageRank ak na ňu odkazuje veľa iných stránok, alebo ak na ňu odkazujú stránky s vysokým PageRankom. PageRank zvláda oba tieto prípady a všetko medzi tým vďaka rekurzívne sa šíriacim váham v odkazovej štruktúre webu. (Brin 2008)

$$PR = 0,15 + 0,85 * (\text{"časť"}^2 \text{ PageRanku každej stránky odkazujúcej na práve hodnotenú stránku})$$

Dá sa povedať, že stránky určujú veľkosť PR pre každú stránku, na ktorú odkazujú. Veľkosť PR, ktorý môžu stránke poskytnúť, je o trochu nižší ako je ich vlastný PR (hodnota ich

---

<sup>2</sup> hodnota PageRanku stránky, z ktorej odkaz prichádza, vydelená počtom všetkých odkazov z danej stránky

vlastného PageRanku \* 0,85). Táto hodnota je potom automaticky rozdelená medzi všetky stránky, na ktoré odkazy smerujú.

Môžeme teda zhrnúť, že odkaz zo stránky s PR4 a 5 odchádzajúcimi odkazmi má väčšiu hodnotu ako odkaz zo stránky s PR8 a 100 odchádzajúcimi odkazmi. PR stránky, ktorá odkazuje na inú stránku je síce dôležitý, no veľkú rolu zohráva aj počet odchádzajúcich odkazov z nej. Čím viac je na stránke odkazov, tým nižšiu hodnotu má PR, ktorý stránkam odovzdáva. Výnimkou sú napríklad odkazové farmy.

Ak by rozdiely medzi hodnotami PR1, PR2, ..., PR10 boli rovnaké, toto zhrnutie by obstálo, no hodnoty medzi PR1 a PR10 (maximálna hodnota, akú môže PageRank dosiahnuť) sú založené skôr na logaritmickú stupnici. Nikto mimo Googlu to nevie naisto, no je veľmi pravdepodobné, že stupnica je logaritmického alebo podobného charakteru. To znamená, že stránka musí vyložiť omnoho väčšie úsilie pre dosiahnutie ďalšieho stupňa PR, ako musela vynaložiť pri získaní predchádzajúceho stupňa. Toto tvrdenie zároveň obracia predchádzajúcu hypotézu tak, že odkaz zo stránky s PR8 a mnohými odchádzajúcimi odkazmi má zrejme väčšiu hodnotu ako zo stránky s PR4 s malým množstvom odkazov.

Ak stránka poskytuje svoj hlas, ktorého hodnota závisí od jej PR, inej stránke, jej vlastný PR sa neznižuje. Nedaruje svoj PR druhej stránke, nejde o presun PR. Je to len hlas, ktorý prepožičiava stránke na základe svojho PR.

Pre výpočet PageRanku určitej stránky, sa jej predchádzajúci PageRank (ak nejaký má) neberie do úvahy a výpočet začína odznova od PR, ktorý stránke prepožičiavajú prichádzajúce odkazy, u ktorých mohla od posledného výpočtu nastať zmena.

#### *3.3.1.4.4 Problém – ukážka na príklade*

Rovnica jasne znázorňuje, ako stránka PageRank získa. Nie je však ihneď zrejmé, že môže fungovať v prípade, že je výpočet vykonaný len raz. Predpokladajme, že máme dve stránky, A a B, ktoré sú navzájom prelinkované a ani jedna z nich nemá žiadne ďalšie odkazy.

**KROK** – výpočet PR stránky A na základe hodnôt jej prichádzajúcich odkazov

Stránka A tak získa novú hodnotu PageRanku. Pri výpočte bola použitá hodnota odkazu zo stránky B, no stránka B má prichádzajúci odkaz od stránky A a jeho nová hodnota PR ešte nebola vypočítaná, teda nový PR stránky A bol vypočítaný na základe nepresnej hodnoty, teda ani on nemôže byť presný.

KROK – výpočet PR stránky B na základe hodnôt jej prichádzajúcich odkazov

Stránka B teraz získala novú hodnotu PR, ktorá však nemôže byť správna, pretože vo výpočte bola použitá nová hodnota PR odkazov zo stránky A, a tá je nepresná. (Craven 2003)

Táto metóda sa zdá byť bludným kruhom. Pre zistenie PageRanku určitej stránky je nutné poznať význam všetkých ostatných stránok, a ten opäť závisí od danej stránky. (Houser 2008)

Ak by sa výpočet zopakoval na základe týchto stanovených hodnôt PR, výsledky by síce boli o niečo presnejšie, no stále by pri výpočte boli použité chybné hodnoty a výsledky by neboli správne. (Craven 2003)

Dá sa tak teda iba overiť, či je určité riešenie správne, teda vyhovuje podmienkam. Úloha však vôbec nemusí mať riešenie, alebo môže mať riešení niekoľko. (Houser 2008)

Problém môže byť prekonaný mnohonásobným opakovaním výpočtu. Každé opakovanie prinesie o trochu presnejšie výsledky. Úplná presnosť ale nemôže byť nikdy dosiahnutá, pretože všetky výpočty sú založené na nesprávnych hodnotách. 40 až 50 opakovaní by malo stačiť pre dosiahnutie bodu, kedy už ďalšie výpočty neprinesú žiadne výrazné zmeny. Tento postup Google vykonáva pri každej aktualizácii a to je dôvod, prečo každá aktualizácia zaberie toľko času.

Treba však brať na vedomie to, že takto získané výsledky vyjadrujú pomer, hodnoty musia byť následne prepočítané podľa stupnice (ktorú pozná len Google) pre získanie PR každej stránky. (Craven 2003)

Google v rámci algoritmu tají aj spôsob, ako z vypočítaného hodnotenia získa verejný PR zobrazovaný v paneloch nástrojov (toolbaroch) alebo na analytických stránkach a postupy, ktoré používa proti podvodníkom, odkazovým farmám a pod. (Houser 2008)

### **3.3.1.5 Alternatívny výpočet**

Niektoré zdroje uvádzajú pre výpočet PageRanku rovnicu:

$$PR(A) = PR(A) + (1-d) + d(\dots)$$

Ide o obdobnú rovnicu už spomenutého výpočtu, do jej výsledku však dopĺňajú existujúci PageRank danej stránky. Ide predovšetkým o situácie, keď je PageRank rozdelený medzi niekoľko stránok.

Výpočet ale môžeme označiť za nesprávny z toho dôvodu, že na základe organizácie odkazov môžu niektoré stránky získať väčšiu časť PageRanku ako ostatné. Pridanie hodnoty súčasného PageRanku aj napriek opakovaniám vyprodukuje iný pomer ako pri použití pôvodného výpočtu. Podľa neho sa kalkulácia pre stránku pri každom opakovaní začína od nuly a spolieha sa výlučne na odkazy, ktoré na ňu smerujú. Myšlienka pridania súčasnej hodnoty PageRanku tento fakt nezahŕňa, a preto nemôže byť správna. (Craven 2003)

### 3.3.1.6 Praktický príklad

Stránka 1 odkazuje na stránku 3, stránka 2 odkazuje na stránku 1, 3 a 4, stránka 3 odkazuje na stránku 2 a stránka 4 neodkazuje nikam.

Vytvorí sa matica, ktorej riadky odpovedajú jednotlivým stránkam a stĺpce odkazom na stránky. Sú v nej zohľadnené váhy odkazov podľa toho, na koľko stránok daná stránka odkazuje. (Houser 2008)

0	0	1	0
1/3	0	1/3	1/3
0	1	0	0
0	0	0	0

Výsledkom sčítania stĺpcov matice sa načrtne váha jednotlivých stránok podľa počtu, resp. kvality, odkazov na ňu.

Stránka 1: 1/3

Stránka 2: 1

Stránka 3: 4/3

Stránka 4: 1/3

Po prepočte budú mať odkazy relatívnu váhu podľa toho, akú váhu ktorá stránka má.

0	0	1/3	0
1/3	0	1/3	1/3
0	4/3	0	0

0	0	0	0
---	---	---	---

Nasleduje opätovné sčítanie stĺpov

Stránka 1: 1/3

Stránka 2: 4/3

Stránka 3: 2/3

Stránka 4: 1/3

A vynásobenie pôvodnej matice novými koeficientmi

0	0	1/3	0
4/9	0	4/9	4/9
0	2/3	0	0
0	0	0	0

Potom by nasledovalo ďalšie spočítanie hodnôt v stĺpcoch.

Matica, s ktorou pracuje Google má však taký počet riadkov, aký je počet všetkých indexovaných stránok.

Problém môžu ale spôsobiť koncové stránky, ktoré už na nič neodkazujú (napríklad stránka č. 4 z príkladu) alebo stránky odkazujúce len samy na seba. PageRank tento problém rieši takým spôsobom, že koncovým objektom pridáva náhodný odkaz.

Po tomto procese nastáva výpočet hodnoty PageRanku a to, akým spôsobom je nakoniec vypočítaná presná hodnota verejného PageRanku zobrazovaná v paneli nástrojov Google, nie je zverejnené, podobne ako ostávajú utajené postupy Google proti podvodným technikám. (Houser 2008)

### 3.3.2 Anchor text

Na rozdiel od iných vyhľadávačov, ktoré text, z ktorého viedol odkaz, hodnotili v rámci stránky, na ktorej sa nachádzal, Google text odkazu priradzoval stránke, na ktorú daný odkaz smeroval. Tento, tzv. anchor text, poskytoval presnejší popis webových stránok ako ony samotné a mohol existovať aj pre dokumenty, ktoré nemohli byť indexované vyhľadávačom

založenom na textoch, ako napríklad obrázky, programy a databázy. Takéto použitie anchor textu teda umožnilo zobrazenie výsledkov oveľa vyššej kvality.

Na druhej strane, nevýhodou bola technická náročnosť kvôli obrovskému množstvu spracovávaných dát, napríklad v prejdenných 24 miliónoch stránok bolo potrebné zaindexovať viac ako 259 miliónov textov odkazov. (Brin 2008)

### **3.3.3 Ďalšie vlastnosti**

Okrem PageRanku a anchor textov používal Google pri hodnotení mnoho ďalších faktorov. Za prvé, keďže mal informácie o umiestnení všetkých hitov, mohol vo veľkej miere využívať vo vyhľadávaní proximitu. Za druhé, Google sledoval niektoré vizuálne detaily prezentácie ako veľkosť fontu slov a slovám napísaným väčším alebo hrubým fontom dával oproti iným slovám väčšiu váhu. A za tretie, úplný zdrojový HTML kód stránok bol pre využitie dostupný z repozitára. (Brin 2008)

## ***3.4 Faktory ovplyvňujúce poradie výsledkov na SERP***

Pri hľadaní určitého hesla sú k dispozícii tisíce až milióny výsledných webových stránok, ktoré danému výrazu zodpovedajú. Z tohto množstva sa vďaka počítačovým programom a vzorcom, ktoré tvoria algoritmus, vyberú tie najrelevantnejšie, ktoré sa užívateľovi následne zobrazia. Tieto hodnotiace algoritmy neustále prechádzajú zmenami a používajú pre hodnotenie stránok viac než 200 faktorov – tzv. signálov. (How Search Works 2012) Tieto signály sa s vývojom a úpravami algoritmu od spustenia vyhľadávača Google zmenili predovšetkým v tom zmysle, že narástol ich počet a zmenila sa u nich miera ich vplyvu na celkové hodnotenie stránky.

Pre lepší prehľad je tieto signály možné rozdeliť podľa ich zamerania do desiatich základných skupín:

### **3.4.1 Faktory týkajúce sa domény**

- vek domény
- kľúčové slovo obsiahnuté v názve domény ideálne na prvom mieste
- dĺžka registrácie domény

- kľúčové slovo v názve subdomény
- história vlastníctva domény
- presná zhoda názvu domény s vyhľadávacím dotazom
- informácie vo WhoIs
- penalizovaný vlastník stránky vo WhoIs

### **3.4.2 Faktory na úrovni stránky**

- Kľúčové slovo v „Title“ tagu
- „Title“ tag začínajúci kľúčovým slovom
- Kľúčové slovo v „Description“ tagu
- Kľúčové slovo v nadpise prvej úrovne (H1 tag)
- Kľúčové slová v najčastejšie používanej vete dokumentu
- Dĺžka textov a rozsah obsahu
- Hustota kľúčových slov (keyword density)
- Kľúčové slová Latentného sémantického indexovania v obsahu
- Rýchlosť načítania stránky cez HTML
- Duplicitný obsah
- Rel=Canonical
- Rýchlosť načítania stránky cez Chrome
- Optimalizácia obrázkov (alt text, nadpis (title), popis (description) a titulok (caption))
- Novosť aktualizácií obsahu (Update Caffeine)
- Dôležitosť updatov obsahu
- História aktualizácií stránky a frekvencia jej updatov
- Hodnota kľúčových slov

- Kľúčové slová v nadpisoch H2, H3
- Poradie slov v rámci kľúčového slova
- Kvalita odchádzajúcich odkazov
- Tematické zameranie odchádzajúcich odkazov
- Gramatika a pravopis
- Syndikovaný obsah
- Užitočný doplnkový obsah (konvertory, kalkulačky, interaktívne návody, ...)
- Počet odkazov zo stránky
- Multimediálny obsah
- Počet vnútorných odkazov smerujúcich na určitú stránku
- Kvalita vnútorných odkazov smerujúcich na určitú stránku
- Zlomené odkazy (Broken Links)
- Úroveň textu
- Odkazy za províziu (Affiliate Links)
- Kvalita kódu (HTML error/WC3 validácia)
- Autorita domény webhostingu stránky
- PageRank stránky
- Dĺžka URL
- Cesta URL (URL path)
- Ľudia v roli editorov
- Kategória stránky
- Wordpress tagy
- Kľúčové slovo v URL



- URL string
- Referencie a zdroje
- Odrážky a číslované zoznamy
- Priorita mapy webu
- Príveľa odchádzajúcich odkazov zahlcujúcich stránku
- Množstvo ďalších kľúčových slov, na základe ktorých je stránka hodnotená
- Vek stránky
- Uživateľsky prijateľné rozloženie stránky
- Parkované domény (Parked domains) (update z decembra roku 2011)

### **3.4.3 Faktory na úrovni webového sídla**

- Hodnota obsahu a unikátneho pohľadu
- Stránka „Kontaktujte nás“
- Dôveryhodnosť domény
- Architektúra webového sídla
- Aktualizácie webu a ich periodicita
- Počet stránok webu
- Mapa webu
- Doba dostupnosti stránky (Site Uptime)
- Lokalizácia servera
- SSL certifikát (weby elektronického obchodu)
- Všeobecné podmienky (Terms of Service) a ochrana súkromia (Privacy Pages)
- Duplicitný obsah webu
- Mininavigácia (Omrvinková navigácia, Breadcrumb Navigation)

- Optimalizácia pre mobilné zariadenia (použitie responzívneho webdesignu)
- Používanie YouTube
- Použitelnosť webu
- Využívanie produktov Google Analytics a Google Webmaster Tools
- Používateľské recenzie / reputácia webu (napr. na Yelp.com alebo na RipOffReport.com)

#### **3.4.4 Faktory súvisiace so spätnými odkazmi**

- Vek odkazujúcej domény
- Počet odkazujúcich domén
- Počet odkazov zo separátnych IP adries triedy C
- Počet odkazujúcich stránok
- Alt tag (pre odkazy z obrázkov)
- Odkazy z domén .edu a .gov
- PageRank odkazujúcich stránok
- Autorita odkazujúcich domén
- Odkazy zo stránok konkurentov
- Zdieľanie odkazujúcej stránky na sociálnych sieťach
- Odkazy od tzv. zlých susedov (Bad Neighborhoods)
- Príspevky návštevníkov
- Odkazy na domovskú stránku
- Nofollow odkazy
- Diverzita typov odkazov
- Výrazy v okolí odkazov

- Odkazy v kontexte
- Nadmerné používanie presmerovania 301
- Anchor texty spätných odkazov
- Anchor texty vnútorných odkazov
- Atribút názvu odkazu
- TLD krajiny odkazujúcej domény
- Umiestnenie odkazu v obsahu stránky
- Relevancia odkazujúcej domény
- Relevancia úrovne stránky
- Sentiment textu v okolí odkazu
- Kľúčové slovo v nadpise
- Rýchlosť rastu počtu odkazov stránky
- Odkazy z top stránok na určitú tému (hub pages)
- Odkazy z tzv. autoritatívnych webových sídiel (Authority Sites)
- Odkaz na web z Wikipédie
- Súvislosti textu v okolí odkazu na určitý web a tematické zameranie tohto webu
- Vek spätného odkazu
- Odkazy z reálnych stránok majú vyššie hodnotenie ako zo splogov
- Prirodzenosť odkazového profilu
- Recipročné (vzájomné) odkazy
- Odkazy v obsahu generovanom užívateľmi
- Odkazy z presmerovania 301
- Používanie mikroformátov

- Umiestnenie v zozname DMOZ
- Umiestnenie v registri Yahoo!
- Počet odkazov smerujúcich zo stránky
- Odkazy z profilov na fórach
- Odkazy z dlhších textov majú vyššiu váhu ako z kratších
- Kvalita obsahu, z ktorého vedie odkaz
- Rovnaké odkazy umiestnené na všetkých stránkach webu

#### **3.4.5 Faktory interakcie používateľov**

- Organický CTR pre určité kľúčové slovo
- Organický CTR pre všetky kľúčové slová
- Miera opustenia
- Návštevnosť z priamych vstupov na stránku
- Opakovaná návštevnosť
- Blokované webové sídla
- Záložky prehliadača Chrome
- Dáta panela nástrojov Google
- Počet komentárov

#### **3.4.6 Špeciálne pravidlá algoritmu**

- Vyhľadávacie požiadavky si zaslúžia čerstvosť
- Vyhľadávacie požiadavky si zaslúžia rozmanitosť
- História prezerania webových stránok používateľom
- História vyhľadávania používateľa
- Geografické cielenie

- Bezpečné vyhľadávanie
- Kruhy (Circles) Google+
- Sťažnosti v rámci DMCA
- Rozmanitosť domén (Aktualizácia Bigfoot)
- Vyhľadávanie kľúčových slov súvisiacich s transakciami
- Lokálne vyhľadávanie
- Okno s Google správami
- Uprednostňovanie veľkých značiek (aktualizácia Vince)
- Výsledky z oblasti nakupovania (Shopping Results)
- Obrazové výsledky (Image Results)
- Jeden výsledok pre doménu každej značky v rámci stránky s výsledkami hľadania

### **3.4.7 Sociálne signály**

- Počet tweetov
- Autorita používateľských účtov na Twitteri
- Počet Facebookových likeov
- Zdieľanie na Facebooku (Share)
- Autorita Facebookových účtov používateľov
- Počet pinov Pinterestu (Pinterest Pins)
- Hlasy na weboch sociálneho zdieľania (ako Reddit, Stumbleupon alebo Digg)
- Počet hodnotení +1
- Autorita používateľských účtov Google+
- Autorstvo overené prostredníctvom Google+ (Google+ Authorship)
- Relevancia sociálnych signálov

- Sociálne signály na úrovni webu

### **3.4.8 Signály týkajúce sa značky**

- Anchor text názvu značky
- Vyhľadávanie názvu značky a súvisiacich výrazov
- Existencia vlastnej stránky značky na Facebooku a jej počet likeov
- Existencia vlastnej stránky značky na Twitter a jej počet Followerov
- Oficiálna stránka spoločnosti na LinkedIn
- Zoznam zamestnancov uvedený na LinkedIn
- Legitimita účtu na sociálnych sieťach
- Zmienky o značke na stránkach s novinkami (News Feed)
- Google berie do úvahy aj zmienky o značke neobsahujúce hypertextový odkaz
- Počet odoberateľov RSS
- Fyzická lokalizácia použitím zoznamu miest Google+
- Google môže overiť, či je určitý web spojený s podnikaním a platením daní

### **3.4.9 Spam na stránke**

- Penalizácia Panda (týkajúca sa stránok s nízkou kvalitou, obsahových fám)
- Odkazy na tzv. zlých susedov
- Presmerovanie
- Pop-Up okná a rušivé reklamy
- Prehnane optimalizované weby
- Prehnane optimalizované stránky
- Príliš reklamných odkazov na stránke (aktualizácia Ads Above the Fold)
- Skrývanie odkazov získaných na princípe dohody (Affiliate Links)

- Sesterské, spolupracujúce weby (Affiliate Sites)
- Automaticky generovaný obsah
- Prehnaná snaha o ovplyvňovanie PageRanku
- IP adresa označená za spam
- Spam v meta tagoch

#### **3.4.10 Spam v okolí stránky**

- Neprirodzený náhly nárast počtov odkazov na stránku
- Penalizácia Penguin
- Odkazový profil s vysokým podielom odkazov nízkej kvality
- Relevancia odkazujúcich domén
- Varovná správa o neprirodzených odkazoch
- Odkazy z rovnakej IP adresy triedy C
- Škodlivý anchor text odkazov smerujúcich na stránku
- Manuálna penalizácia (Manual Penalty)
- Predaj odkazov
- Umiestnenie do Google Sandbox
- Google Dance
- Nástroj na oznámenie odkazov, ktoré by nemali byť uznané (Disavow Tool)
- Žiadosť o znovuposúdenie (Dean 2013)

## 4 UKÁŽKA HODNOTIACICH SIGNÁLOV NA PRAKTICKOM PRÍKLADE

### 4.1 Úvod do výskumu

Hodnotenie webových stránok, na základe ktorého sú následne usporadúvané na stránke s výsledkami hľadania, je zložitým procesom založeným na analýze obrovskej skupiny signálov. Tento proces je základom jedinečnosti vyhľadávača Google a preto spoločnosť jeho konkrétnu podstatu prísne tají. Na základe priebežných oficiálnych vyhlásení a pozorovaní rankingov, a ich zmien sa však podarilo vytvoriť zoznam faktorov, u ktorých sa predpokladá, že sú vo väčšej alebo menšej miere súčasťou hodnotiaceho algoritmu. V priebehu času a vývoja vyhľadávača sa hodnota spomenutých signálov mení – v dnešnej dobe je už vplyv niektorých faktorov, ktoré boli v počiatkoch fungovania vyhľadávača Google nosnými kritériami hodnotenia, takmer nepatrný, a na druhej strane pribudol veľký počet nových signálov.

Najviac významných zmien sa objavilo v posledných rokoch, kedy sa do popredia stále viac dostávajú sociálne signály a s rapídny vývojom sociálnych sietí ich počet rýchlo stúpa. Predchádzajúca kapitola obsahuje náčrt signálov z niekoľkých oblastí, ktoré boli zaznamenané na základe pozorovania počas celého vývoja Google. Rozdelené sú podľa ich povahy na desať skupín – faktory na úrovni domény, faktory na úrovni stránky, faktory na úrovni webového sídla, faktory súvisiace so spätnými odkazmi, faktory interakcie používateľov, špeciálne pravidlá algoritmu, sociálne signály, signály týkajúce sa značky, spam na stránke a spam v okolí stránky. Niektoré z nich budú v tejto kapitole ukázané na vzorovej skupine náhodne vybraných konkrétnych výsledkoch vyhľadávania.

### 4.2 Výber kľúčového slova

Na začiatku výskumu bolo potrebné správne zvoliť kľúčové slovo, pre ktoré by bolo možné z výsledkov jeho vyhľadávania vybrať niekoľko odkazov pre analýzu hodnotiacich signálov.

Predpokladaný typ odkazov pre výskum musel spĺňať niekoľko podmienok:

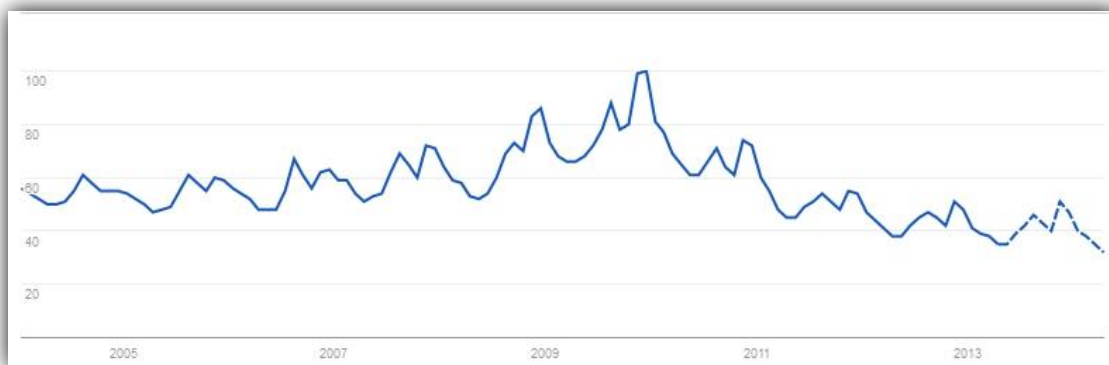
- *Konkurenčné prostredie* – nabáda k neustálemu zlepšovaniu webu a optimalizácii stránky



- *Využívanie sociálnych sietí* – potrebné pre analýzu vplyvu sociálnych signálov
- *Dĺžka existencie webu* – u nových webov stránky ešte nemuseli prejsť procesom indexácie a hodnotenia, nemajú históriu ani dostatok faktorov hodnotenia
- *Odkazový profil* – odkazy sú jedným z najpodstatnejších kritérií, preto bolo nevyhnutné zamerať sa na stránky s potenciálom čo najrozvinutejšej odkazovej štruktúry
- *Zameranie na technickú stránku* – weby, na ktorých je možné hodnotiť technické prvky
- *Rýchle sa vyvíjajúca oblasť* – stránky z oblasti, v ktorej je možné predpokladať využívanie najnovších trendov v oblasti internetu
- *Vysoká hľadanosť* – výber kľúčového slova, ktoré je vyhľadávané a podporuje konkurenčné prostredie
- *Známosť značky* – jednu skupinu hodnotiacich faktorov tvoria signály týkajúce sa značky, odkazy pre analýzu mali preto patriť známym spoločnostiam, aby sa na nich tieto signály dali pozorovať

Po zvážení cieľu výskumu a dôvodov pre výber odkazov, a na základe mnohých pokusov a vyhľadávaní, bolo pre analýzu zvolené kľúčové slovo „laptops“. Množné číslo je použité z dôvodu konkrétnejšieho obmedzenia výsledkov a na základe porovnania povahy výsledkov nájdených pri hľadaní výrazov „laptop“ a „laptops“. Na kľúčové slovo „laptop“ boli na prvých pozíciách prevažne stránky obsahujúce recenzie produktov, ale na kľúčové slovo „laptops“ už výsledky na prvej stránke zahŕňali stránky konkrétnych konkurenčných spoločností – výrobcov týchto produktov.

Na obrázku č. 3 vidíme zmeny v počte hľadání tohto kľúčového slova od roku 2004 – vrchol bol zaznamenaný v roku 2010 a aj napriek tomu, že za posledné roku táto hľadanosť oproti minulému desaťročiu klesla, objemy hľadání sa pohybujú v číslach postačujúcich pre analýzu. Zároveň odhadovaná predpoveď vyhľadávania tohto kľúčového slova (v obrázku vyznačená prerušovanou čiarou) ukazuje zachovanie týchto objemov i v najbližšom období.



Obr. č. 3: Vývoj hľadania výrazu „laptops“ od roku 2004 po súčasnosť

(Zdroj: <http://www.google.com/trends/>).

### 4.3 Výber odkazov

Výber odkazov, ktoré boli následne použité v analýze, bol zameraný na prvú stránku výsledkov hľadania, teda na prvých desať odkazov zobrazených na vopred zvolené kľúčové slovo (Obr. č. 4). V ďalšom kroku sme separovali 5 výsledkov a to:

- a) prvý v poradí – elektronický obchod
- b) stránky navzájom si konkurujúcich spoločností vyrábajúcich laptopy
- c) stránku časopisu, ktorý obsahuje informácie a recenzie o týchto výrobkoch

Vybrané odkazy sa nachádzali na pozíciách 1, 3, 6, 8 a 10, čím sa nám podarilo obsiahnuť skupinu najrelevantnejších výsledkov pri hľadaní kľúčového slova „laptops“. Do úvahy sme nebrali platené výsledky ani vyhľadávanie v správach, pretože výskum bol zameraný na organické výsledky a spôsob hodnotenia ich relevancie.

Na obrázku č. 4 sú na stránke s výsledkami vyhľadávania po zadaní hľadaného výrazu „laptops“ znázornené vybrané odkazy na analýzu.



Obr. č. 4: Stránka výsledkov vyhľadávača Google pri hľadaní hesla „laptops“

(Zdroj: <http://www.google.com>)

## 4.4 Výber hodnotiacich faktorov

Z početnej skupiny faktorov v kapitole číslo 3 sme pre analýzu vybrali tie faktory, ktoré by mohli byť z pohľadu výskumu najzaujímavejšie, hoci pre nás najobmedzujúcejším faktorom bola dostupnosť dát. Ideálne hodnotenie najpodstatnejších faktorov bolo z tohto dôvodu možné len čiastočne. K dispozícii sme mali prevažne online nástroje a niekoľko platených aplikácií, nemali sme však prístup do štatistík týchto webov, a pre mnohé signály neexistuje dostupný nástroj, ktorý by ich vedel merať.

Treba však zdôrazniť, že čísla nemusia presne zodpovedať realite, a preto je potrebné brať ich len ako orientačné hodnoty, na základe ktorých je možné porovnať stav meraných webových stránok. Na základe použitia niekoľkých nástrojov pre každý hodnotený signál sme overili spoľahlivosť pomeru týchto hodnôt.

## 4.5 Analýza

Analyzované stránky a ich domény sme si pre lepšiu prehľadnosť označili veľkými písmenami (tabuľka č. 1)

Webová stránka s URL adresou <http://www.bestbuy.com/site/Computers-PCs/Laptop-Computers/abcat0502000.c?id=abcat0502000> a doménou *bestbuy.com* má označenie A, stránka <http://www.dell.com/us/p/laptops> na doméne *dell.com* označenie B, stránka <http://www.lenovo.com/products/us/laptop/> na doméne *lenovo.com* označenie C, stránka <http://www.pcworld.com/category/laptop-computers> na doméne *pcworld.com* označenie D a stránka s URL <http://www.hp.com/country/us/en/prodserv/laptops.html> a doménou *hp.com* dostala označenie E.

Tabuľky č. 1 – 11 v prílohe obsahujú hodnoty parametrov pre tieto stránky rozdelené podľa skupín hodnotiacich kritérií definovaných v kapitole č. 3 tejto práce. Tabuľky neobsahujú kritéria skupiny faktorov interakcie používateľov (ich dostupnosť je limitovaná prístupom k štatistikám stránok) a skupín faktorov súvisiacich so spamom (tieto faktory znižujú kvalitu stránky a následne jej hodnotenie, vo vážnych prípadoch je web penalizovaný, čo sa pozorovaných webov v tomto prípade netýka).

### 4.5.1 Analýza faktorov na úrovni domény

Z databázy WhoIs sme získali presné dátumy vytvorenia domény, jej veku, dátum skončenia jej platnosti a dátum poslednej aktualizácie záznamu (tabuľka č. 2). Tieto faktory majú údajne

minimálny až takmer žiadny vplyv na hodnotenie stránok, v tomto prípade však mala najlepšie postavenie doména hp.com, ktorá je vo veku 27 rokov a 2 mesiace najstaršou spomedzi ostatných domén, a čím je doména staršia, tým sa predpokladá jej vyššia dôveryhodnosť. Pri hodnotení dátumu, do ktorého je doména platná, sa domény zaplatené na niekoľko rokov dopredu považujú za právoplatnejšie a spoľahlivejšie, čo im môže pridať na kladnom hodnotení. Z tohto pohľadu by mala najlepšie postavenie doména pcworld.com a hneď za ňou doména bestbuy.com.

Keď hodnotíme prítomnosť skúmaného kľúčového slova „laptops“ v názve domény, čo je jeden zo závažnejších faktorov, všetky monitorované domény získali rovnaký výsledok – pretože na toto kľúčové slovo sú zamerané a optimalizované len ich podsekcie a nie je to ich hlavné tematické zameranie, v názve domény sa slovo „laptops“ ani v jednom prípade nenachádza.

#### **4.5.2 Analýza faktorov na úrovni stránky**

Z tejto skupiny hodnotiacich signálov bolo pre analýzu dostupných najviac dát (tabuľky č. 3 a č. 4). Pri hodnotení výskytu nami zvoleného kľúčového slova v meta tagoch skúmaných stránok získala najlepšie hodnotenie stránka D, ktorá v nadpise obsahuje zvolené kľúčové slovo na prvom mieste a popise stránka sa objavuje až trikrát. Druhá najúspešnejšia bola stránka E. Keď však zhodnotíme len nadpisy stránok, všetky stránky, okrem stránky C, majú kľúčové slovo „laptop“ na 1. mieste, čo ich stavia do rovnakej pozície, rovnako ako pri hodnotení nadpisov najvyššieho radu, kde žiadna z analyzovaných stránok nezaostáva.

Dĺžka obsahu stránky zrejme na hodnotenie výrazný vplyv nemá a žiadna z týchto stránok výrazne nevyčnieva. Ďalej sme hodnotili percento výskytu slova „laptops“ medzi ostatnými kľúčovými slovami na týchto stránkach. Najvyšší výskyt má u stránky E (1,87%), u ostatných sa pomer pohybuje medzi 1,1% a 1,4%. Rýchlosť načítania stránky Google hodnotí tzv. Google score, na základe ktorého mala najlepší výsledok stránka D, ďalej stránka B, C, E a nakoniec stránka A. Obrázky sú najlepšie označené na stránke B, stránka C nepoužíva nadpis ani alt text. Tieto označenia obrázkov sú nápomocné nielen pri hľadaní v Google obrázkoch, ale alt text obrázkov plní funkciu anchor textu, v prípade, že z nich vedie odkaz.

Hodnota autority stránky, ďalší relatívne významný faktor, je vypočítaná cez celkovú hodnotu autority jej odkazov a z analyzovaných stránok ju má najvyššiu stránka B, hneď za ňou stránka A a ostatné stránky. Zlomené odkazy sú odkazy nefunkčné a Google môže pri ich väčšom množstve považovať stránku za nepoužívanú. Zo skúmanej skupiny domén

má najviac zlomených odkazov jednoznačne stránka E, či už pri hodnotení počtu týchto odkazov alebo percentuálneho zastúpenia medzi všetkými odkazmi stránky. Najviac stránok s chybami v zdrojovom kóde má web pcworld.com a najmenej bestbuy.com. Tento faktor nepatrí k najvplyvnejším, no veľký počet chýb v kóde môže mať negatívny dopad.

PageRank stránky, ktorému je v tejto práci venovaná celá kapitola, je jedným z najviac vplyvných faktorov určujúcich poradie zobrazených stránok na určité kľúčové slovo. PageRank stránky býva spravidla vyšší ako PageRank celého webu, z monitorovaných stránok majú PR 6 stránky na najvyšších pozíciách – A a B, stránka C má PR 5, u stránky D ho nebolo možné zistiť a stránka E má PR 0, čo môže byť podľa Šubra (2013) spôsobené napríklad dočasným problémom spôsobeným prepočtom algoritmu, penalizáciou Google alebo ďalšími vplyvmi. Kľúčové slovo v URL stránky má zrejme tiež určitú hodnotu a z tohto pohľadu majú výhodné postavenie stránka B a stránka E. Mapu webu využívajú všetky skúmané stránky a okrem toho, že pomáha robotom pri indexovaní stránky, predstavuje aj jej štruktúru, a v rámci nej postavenie a váhu jednotlivých stránok.

#### **4.5.3 Analýza faktorov na úrovni webového sídla**

V tretej hodnotenej oblasti faktorov (tabuľka č. 5) určujúcich hodnotu webu sa monitorované domény zhodujú v niekoľkých bodoch – majú rovnakú lokalizáciu us (Spojené Štáty) a všetky obsahujú všeobecné podmienky, podmienky ochrany súkromia a stránku „Kontaktujte nás“ s kontaktnými údajmi spoločnosti, čo im môže zvýšiť dôveryhodnosť. Tú Google hodnotí prostredníctvom tzv. TrustRanku, ktorý mu pomáha odlišovať kvalitný obsah od spamu. Najvyšší TrustRank, teda dôveryhodnosť, majú weby dell.com a hp.com, tesne za nimi nasleduje web bestbuy.com, potom lenovo.com a na poslednom mieste je s veľkým odstupom a TrustRankom len 4,65 bodov z 10 web pcworld.com.

Ani pri jednej z nich nebol zistený duplicitný obsah, ktorý by z pohľadu hodnotenia stránok mohol mať negatívny vplyv. Počet stránok Google indexu v tabuľke číslo 5 pochádza z nástroja Website Auditor a vzhľadom na obrovské rozdiely medzi hodnotami pre nás táto informácia nie je príliš spoľahlivá. Orientačne by sa dalo zhrnúť, že najviac indexovaných stránok majú webové stránky na vyšších pozíciách, teda stránky A a B.

Kanál na portáli YouTube majú všetky monitorované weby, najviac odberateľov má kanál webu lenovo.com a hneď za ním dell.com. Keďže sa videá z YouTube vo výsledkoch vyhľadávania objavujú na preferovaných pozíciách, ich vkladanie alebo embedovanie na webové stránky môže dopomôcť k získaniu vyšších pozícií.

Štatistiky webov sú cez nástroj Google Analytics merané len u webu pcworld.cz, ostatné weby používajú vlastné nástroje.

#### **4.5.4 Analýza faktorov súvisiacich so spätnými odkazmi**

Odkazový profil každého webu je bezpochyby jedným z najhlavnejších metrík používaných pre určovanie výsledkov vyhľadávania. Ide predovšetkým o počet odkazov, bez ohľadu na to, či hodnotíme počet odkazujúcich, domén, počet odkazujúcich stránok, počet odkazov zo separátnych IP adries triedy C alebo počet IP adries smerujúcich na web.

Tabuľky č. 6 a 7 obsahujú niekoľko čísel pre porovnanie počtu odkazov na web, na domovskú stránku webu alebo len na monitorovanú stránku. Pre globálny náhľad na odkazový profil vybraných webových sídiel sa zameriame na celkový počet domén smerujúcich na web, pričom weby môžeme zoradiť v nasledujúcom poradí: E, B, D, A a C. Pri počte odkazujúcich domén len na vybranú stránku je však poradie odlišné: B, A, D, C a E. Z toho sa dá predpokladať, že Google uprednostňuje počet odkazov na konkrétnu stránku, ktorú vo výsledkoch zobrazí, pred počtom odkazov na web, hoci o dôležitosti ani jednej z týchto metrík nemôže byť pochyb.

Niektorí správcovia webov dávajú dôraz aj na doménu prichádzajúcich odkazov a uprednostňujú napríklad domény .edu a .gov, čo podľa Matta Cuttsa údajne nemá na hodnotenie žiaden vplyv.

Významným prvkom hodnotenia odkazov prichádzajúcich na stránku je aj ich anchor text. V tabuľke č. 7 vidíme 3 najčastejšie používané anchor texty pre každý z analyzovaných webov – vo väčšine prípadov sa v nich objavuje názov značky, prípadne neobsahujú žiaden text. Rýchlosť rastu počtu odkazov na web je možné odhadnúť z grafov na obrázkoch č. 6 – 10 v prílohe. Vidíme, že z dlhodobého hľadiska u všetkých webov počet odkazov klesol, hoci v poslednom období začína opäť rásť (okrem webu pcworld.com). U žiadneho z nich sa ale neobjavili príliš výrazné zmeny vo vývoji počtu odkazov. Presmerovanie odkazov môžu znížiť ich PageRank tým, že ho medzi nich rozdelí. Najhoršiu pozíciu z tohto pohľadu má web hp.com a lenovo.com, na druhej strane, žiaden odkaz z presmerovania nebol nájdený pri webe bestbuy.com.

Určitý prínos v rámci hodnotenia stránok môže mať aj umiestnenie v zozname DMOZ a v registri Yahoo!, čo spĺňajú všetky sledované weby s výnimkou lenovo.com, ktorý sa v registri Yahoo! nenachádza.

#### **4.5.5 Analýza špeciálnych pravidiel algoritmu**

Sociálna sieť spoločnosti Google nazvaná Google plus sa stala z jedným hodnotiacich faktorov webových stránok. Z monitorovaných webov má každý vlastnú firemnú stránku Google+ (tabuľka č. 9), jej význam zvyšuje počet ľudí, ktorí si ju pridali do svojich kruhov. Najviac ľudí má v kruhoch stránku spoločnosti Dell – viac než 300 tisíc, na druhom mieste je stránka Lenovo, tesne za ňou stránka HP, najmenej ľudí má v kruhoch stránka Bestbuy spolu so stránkou Peworld. Všetkých 5 webov sa objavuje vo výsledkoch hľadania v Google správach, a taktiež vo výsledkoch vyhľadávania z oblasti nakupovania. Pri zadaní kľúčového slova „laptops“ pri vyhľadávaní v obrázkoch sa medzi prvými dvadsiatimi výsledkami objavili po dva výsledky pre weby bestbuy.com a pcworld.com a jeden výsledok pre dell.com.

#### **4.5.6 Analýza sociálnych signálov**

Signály zo sociálnych sietí naberajú v posledných rokoch stále viac na dôležitosti a majú stále významnejší vplyv na ranking webových stránok (tabuľka č. 10). Z analyzovaných webov je vzhľadom na počet tweetov, likeov a zdieľaní obsahu na Facebooku najúspešnejší web bestbuy.com. Na druhom mieste nasleduje web dell.com a na treťom hp.com. Na Pintereste je najúspešnejší s prevahou web hp.com. V počte hodnotení +1 v rámci siete Google Plus môžeme pozorované weby zoradiť podľa úspešnosti takto: hp.com, dell.com, lenovo.com, bestbuy.com a nakoniec pcworld.com. Autorstvo na svojich stránkach používa len značka Lenovo.

#### **4.5.7 Analýza signálov týkajúcich sa značky**

Najvyhľadávanejší z vybraných názvov webov (značiek) je „dell“, za ním nasleduje „best buy“, „lenovo“, „pc world“. A na poslednom mieste je „hp“. Objemy hľadání sa však v čase menia a aj údaje obsiahnuté v tabuľke č. 11 považujeme len za orientačné.

Z vybraných stránok v rámci sledovaného konkurenčného prostredia má každá svoje oficiálne stránka na najpoužívanejších sociálnych sieťach ako je Facebook, Twitter, LinkedIn, jedinou výnimkou je neexistujúca stránka spoločnosti PC world na LinkedIn. Keď ich hodnotíme s ohľadom na komunitu, ktorú si na týchto sieťach vybudovali, každá značka dominuje inej sieti. Najviac priaznivcov na Facebooku má značka Best Buy, Twitteru dominuje PC World a profil spoločnosti s najväčším počtom zamestnancov má spoločnosť HP. Všetky spoločnosti je možné lokalizovať prostredníctvom Google miest.



#### 4.5.8 Vyhodnotenie

Pri vyhľadávaní kľúčového slova „laptops“ sa na prvej pozícii výsledkov hľadania objavila stránka s URL adresou <http://www.bestbuy.com/site/Computers-PCs/Laptop-Computers/abcat0502000.c?id=abcat0502000>. Keď sa pozrieme na faktory zaznamenané v tabuľkách č. 2 – č. 11 v prílohe, najlepšie výsledky spomedzi ostatných analyzovaných stránok (webov) dosiahla u týchto faktorov:

- Dĺžka platnosti domény
- Kľúčové slovo v nadpise stránky a v nadpise H1
- Najviac obsahu na stránke s výrazným výskytom vybraného kľúčového slova
- Najmenej chýb v zdrojovom kóde
- Druhá najvyššia autorita stránky
- PR stránky je 6
- Obsahuje vybrané kľúčové slovo „laptops“ v URL adrese
- Druhý najvyšší TrustRank
- Veľmi vysoký počet zaindexovaných stránok (vyšší ako väčšina sledovaných webov)
- Žiadne odkazy z presmerovania
- Dobrá optimalizácia obrázkov – na kľúčové slovo „laptops“ sa vo vyhľadávaní obrázkov zobrazila dvakrát v top 20 výsledkoch
- Prvenstvá v sociálnej oblasti: najvyšší počet tweetov, likeov a zdieľaní na Facebooku
- Veľký mesačný objem hľadání názvu značky
- Facebooková stránka spoločnosti s obrovským počtom fanúšikov

Aj napriek tomu, že tento web v počte odkazov nepatril k najúspešnejším a nemal ani najvyšší PageRank, umiestnil sa na prvom mieste. Tento fakt dokazuje, že spôsob, akým Google stránky hodnotí, je založený na zložitej kombinácii obrovského počtu signálov a nie je možné určiť jeho jednoznačný postup. Sledovanie týchto potenciálnych vplyvov však má svoj význam – naznačuje tvorcom webových stránok akým smerom sa majú uberať, a je dôkazom

neustálych zmien hodnotiaceho algoritmu za účelom dosiahnutia presnejších výsledkov na určitú hľadajú požiadavku.

Analýza zahŕňala len zlomok faktorov, ktoré určujú hodnotu stránky, z výsledkov je však zreteľný význam kvality stránky, jej obsahu a sociálnych faktorov. Na prvom mieste v hodnotení už nestoja primárne odkazy, ale dobiehajú ich signály zo sociálnych sietí – hodnota webových stránok je čoraz viac závislá na bežných používateľoch. Svoj názor na obsah môžu vyjadriť rôznymi funkciami týchto sietí a zdieľaním obsahu jedným kliknutím sa výrazne zvyšuje potenciál jeho šírenia.

Významnou skupinou sú aj faktory, ktoré ovplyvňujú dôveryhodnosť stránky. Tá sa dá zvýšiť transparentným prístupom nielen v rámci budovania odkazového profilu ale aj kvalitným a aktuálnym obsahom bez používania zakázaných praktík, ako je napríklad skrytý alebo príliš malý text, cloaking, odkazové farmy, prepĺňanie stránky kľúčovými slovami, klamné presmerovanie stránok, doorway pages, duplicitný web alebo stránky, tapetovanie katalógov, cybersquatting, keyword stuffing a ďalšie. Pri odhalení používania takýchto metód hrozí stránke penalizácia – zhoršenie jej hodnotenia a zníženie rankingu, v obzvlášť závažných prípadoch môže byť stránka vyradená z indexu, a tým pádom sa prestane objavovať na stránke s výsledkami hľadania.

## 5 NAJVÝZNAMNEJŠIE ZMENY ALGORITMU

### 5.1 Aktualizácie v rokoch 1998 - 2002

Google mení svoj algoritmus vyhľadávania 500-600 krát za rok. Väčšina z týchto zmien je minimálnych, no raz za niekoľko mesiacov Google vykoná väčšiu zmenu algoritmu, ktorá sa prejaví výraznými rozdielmi vo výsledkoch vyhľadávania.

Všetko začalo v marci roku 1998, kedy Google ako prvý spustil vyhľadávač fungujúci na princípe hodnotenia a usporiadania výsledkov s ohľadom na odkazy. O necelé tri roky neskôr, dňa 11. decembra roku 2000, bolo v Mountain View v Kalifornii vyhlásené oficiálne spustenie panela nástrojov (Toolbar, TB). Tento voľne dostupný zásuvný model pre vyhľadávače umožnil užívateľom internetu vyhľadávať vysoko relevantné informácie rýchlo z ktorejkoľvek stránky na internete. Poskytoval možnosť vyhľadávania informácií bez hlbších rešeršných schopností a rýchle vyhľadávanie na ktorejkoľvek stránke.

Medzi jeho ďalšie výhody patrilo:

- Vyhľadávanie stránok akéhokoľvek webového sídla na internete, aj keď dané sídlo nemá svoj vlastný vyhľadávač
- Automatická lokalizácia hľadaných výrazov na ktorejkoľvek webovej stránke, bez použitia funkcie "Hľadať" (Find)
- Označenie každého z termínov vlastnou farbou
- Priamy prístup k webovým stránkam uloženým vo vyrovnávacej pamäti Google (cache) v prípade, že daný odkaz nefunguje alebo ak bol prevzatý offline
- Vyhľadávanie príbuzných stránok
- Automatické určovanie dôležitosti pre každú prehliadanú stránku prostredníctvom Google PageRank technológie
- Vyhľadávanie vo viac než 1,3 miliónoch webových stránok z akéhokoľvek sídla na internete (Google Launches The Google Toolbar 2012)

Zároveň bol vyvinutý aj jeho hodnotiaci faktor – Toolbar PageRank (TBPR), a tým začala éra optimalizácie pre vyhľadávače (Search Engine Optimization, SEO).

Niekoľko prvých zmien v hodnotení stránok bolo len zmenou kombinácií algoritmov a prebiehali na mesačnej aktualizácii indexu nazývanej aj Google Dance.

Prvá zdokumentovaná zmena algoritmu nastala v septembri roku 2002 a hoci jej podstata bola nejasná, bolo isté, že ide o niečo viac než o bežnú aktualizáciu PageRanku. (Gupta 2012)

Rozruch spôsobilo objavenie sa stránok 404 medzi prvými desiatimi výsledkami vyhľadávania a zdalo sa, že PageRanku hrozí zánik. (Google Dance Infographic finished 2011)

## **5.2 Aktualizácie roku 2003**

### **5.2.1 Boston**

O 5 mesiacov po tejto zmene, vo februári 2003 (Gupta 2012), bola v centre doplnkového vzdelávania v Bostone na konferencii Search Engines Strategies (Stratégie vyhľadávačov) (Supplemental education services, SES) vyhlásená prvá oficiálna aktualizácia Google, ktorá získala pomenovanie Boston. (Google Algorithm Change History 2012)

Zameraná bola na odkazy, ktoré na stránku prichádzajú, a na unikátny obsah. (Google Updates Log 2011) Táto aktualizácia odstránila predovšetkým hodnotu, ktorú Google pripisoval odkazom z domén, ktorým už vypršala platnosť. Okrem toho apelovala aj na hodnotu odkazov z registrov, ktorých prínosom si Google nebol celkom istý. (Google Updates Through The Years 2011) Výsledkom bolo, že mnohí správcovia webových sídiel ohlásili straty spätných odkazov a s tým súvisiace zníženie hodnoty PR. (Google Updates Log 2011)

Pôvodne bola cieľom Google veľká mesačná aktualizácia, takže niekoľko prvých vylepšení bolo kombináciou zmien v algoritme s obnovením indexu. Keď sa aktualizácie stali častejšími, myšlienka pravidelného mesačného aktualizovania zanikla. (Google Algorithm Change History 2012)

### **5.2.2 Cassandra**

Onedlho na to bola ohlásená zmena algoritmu Cassandra, ktorá okrem základných otázok kvality odkazov zahŕňala aj útok na skryté odkazy a skrytý text. (Gupta 2012)

Google v rámci tejto aktualizácie zakročil aj v niekoľkých základných otázkach kvality prepojenia, napríklad masívneho prepojenia domén spoluvlastníkov. (Google Algorithm

Change History 2012) Hlavnou myšlienkou bolo, aby si spoločnosti zvolili taký názov, ktorý by odrážal ich názov domény. (Google Updates Log 2011)

### **5.2.3 Dominic**

Zmena algoritmu známa aj ako "Your guess is as good as mine update" – Vaša domnienka je tak dobrá ako moja aktualizácia. (Google Updates Through The Years 2011)

V máji 2003 Google vykonal v algoritme vyhľadávania ďalšie zmeny v rámci (Gupta 2012) aktualizácie Dominic, ktorá bola pomenovaná po jednej pizzerii v Bostone, často navštevovanej účastníkmi konferencie PubCon. (Google Updates Log 2011) Hoci sa máji 2003 udialo mnoho zmien, skutočná povaha algoritmu Dominic bola veľmi nejednoznačná. (Google Algorithm Change History 2012) experti na oblasť optimalizácie webových stránok si určité zmeny všimli, no aj napriek tomu táto pomerne veľká aktualizácia nebola v skutočnosti nikdy úplne vysvetlená, ignorujúc tisícky príspevkov na fórach webmasterov. (Google Updates Through The Years 2011)

Predpokladá sa, že bola zameraná najmä na vytvorenie procesu tematického vyhľadávania a prepojenie datacentra s príslušným vyhľadávaním. Aktualizácia Dominic objasnila zámer orientácie každého datacentra na iný účel. (Google Updates Log 2011)

V tomto období sa stal nástroj Google Dance skutočným fenoménom, keď roboty nazvané „Freshbot“ a „Deepcrawler“ navštívili webové sídla WWW. (Google Updates Through The Years 2011) Po prehľadaní webu mnoho lokalít hlásilo odchýlky. Zdalo sa, že spôsob, akým Google počíta a eviduje spätné odkazy, čakajú radikálne zmeny. (Google Algorithm Change History 2012)

### **5.2.4 Esmeralda**

Tretou aktualizáciou v poradí bola aktualizácia Esmeralda, ktorá uprednostňovala stránky obsahujúce špecifickejšie informácie pre návštevníkov. Odhalila, že aktualizácia Dominic mohla interným stránkam vnútri webových sídiel priradovať vyššiu relevanciu, čo zrejme zvýhodnilo domovskú stránku oproti vyhľadávaniam špecifických dotazov. Užívatelia potom začali oznamovať podstatne menej spamu ako po aktualizáciách Dominic a Cassandra. (Google Updates Log 2011)

Esmeralda zároveň predstavuje ukončenie pravidelných mesačných aktualizácií a Google prechádza sa na kontinuálnejší proces aktualizácie. Google Dance nahradí Everflux a Esmeralda je akousi predzvesťou zásadných zmien Google infraštruktúry.

#### **5.2.5 Fritz**

Definitívny koniec pravidelných mesačných aktualizácií Google Dance nastal až s aktualizáciou Fritz, ktorá bola vyhlásená v júli roku 2003. Namiesto kompletnej úpravy indexu na mesačnej báze zvolil Google inkrementálny prístup a index sa začal meniť takmer každý deň. (Google Algorithm Change History 2012)

Keďže boli aktualizácie pomerne časté, mnohé z nich boli len malými dennodennými zmenami, ktoré sa začali označovať výrazom Evelflux (neustály tok). (Cutts 2006)

#### **5.2.6 Supplemental index**

V septembri 2003 oddelil Google niektoré výsledky do tzv. Doplnkového indexu (Supplemental Index), aby mohol indexovať väčšie množstvo dokumentov bez obmedzenia výkonu. To sa však stalo horúcou témou v oblasti optimalizácie webových stránok pre vyhľadávače, a nakoniec bol Doplnkový index znovu začlenený do hlavného indexu. (Google Algorithm Change History 2012)

#### **5.2.7 Florida**

Aktualizácia Florida v novembri 2003 ukazuje posun Googlu od jednoduchých filtrov k pokusu o pochopenie obsahu vo vzťahu ku kontextu vyhľadávania a potenciálne vyhladaných výsledkov. Florida odstránila spam pomocou jednoduchého spojenia odkazov s ďalšími funkciami, čo dalo väčší význam dobre optimalizovaným stránkam s čistými odkazmi. Správcovia webov aktualizáciu uvítali, čo ukázalo, že hlavnou prioritou Googlu sú záujmy spotrebiteľov. Aktualizácia Florida bola pokusom o podporu stránok, ktoré dodržiavali požiadavky na kvalitu. (Google Updates Log 2011)

Aktualizácia Florida pohla celým SEO priemyslom – mnoho stránok stratilo svoj ranking a ich vlastníci zúril. (Google Algorithm Change History 2012) Predtým vysoko hodnotené stránky prestali bez jasného dôvodu v indexe existovať, čo spôsobilo pokles faktora relevancie a oslabilo tematické väzby. (Google Dance Infographic finished 2011)

Florida znamenala koniec SEO praktík 90. rokov s minimálnou hodnotou, ako je napríklad prepĺňanie stránok kľúčovými slovami, a SEO oblasť začala byť oveľa zaujímavejšia. (Google Algorithm Change History 2012)

### **5.3 Aktualizácie roku 2004**

#### **5.3.1 Austin**

Po aktualizácii Florida prišla v januári 2004 situáciu napraviť aktualizácia Austin. Google sa aj v tomto prípade stále zameriaval na podvodné techniky uplatňované na stránkach, ako je neviditeľný text a prepĺňovanie stránok meta tagmi. (Google Algorithm Change History 2012)

Update sa dotkol aj lokálneho ranku, LSI, automatického skracovania slov podľa ich koreňov (stemming) a iných v tejto dobe používaných optimalizačných techník. (Google Dance Infographic finished 2011)

Hlavným terčom pozornosti sa však stal postup zvaný bombardovanie Google (Google Bombing), ktorý spočíval v tom, že ľudia manipulovali so systémom tak, že ukazoval skreslené výsledky. Pozornosť sa presunula na webové sídla s minimálnou hustotou kľúčových slov a s kvalitnými vnútornými odkazmi. Relevantným odkazom bola pridelená väčšia váha, ak dané webové sídlo bolo odkazmi prepojené so sídlami podobne zameranými, a tým mali vo výsledkoch vyhľadávania lepšie postavenie. (Google Updates Log 2011)

Hovorí sa, že v tomto období dosiahol Google určitú hranicu a začal brať vážne relevanciu indexovaných stránok. (Google Algorithm Change History 2012)

#### **5.3.2 Brandy (známa aj ako Vince)**

Aktualizácia Brandy vo februári 2004 síce nemala vplyv na väčšinu otázok, no zmenila štruktúru zoznamu výsledkov vyhľadávania v prípadoch, keď sa v požiadavke vyskytol názov obchodnej značky. Napríklad pri zadaní "Mustang" do vyhľadávania sa zobrazilo aj množstvo príbuzných webov o autách. Po tejto aktualizácii teda mohli byť na daný výraz zobrazené rôzne stránky súvisiace s autami, koňmi alebo aj športovými tímami, ak mali tento výraz v názve, a podobne. (Google Updates Through The Years 2011)

V rámci tejto aktualizácie teda Google podstatne rozšíril svoj index, zaviedol latentné sémantické indexovanie a zvýšenú pozornosť venoval relevancii anchor textov a konceptu príbuzných odkazov. LSI zvýšilo vyhľadávaču Google schopnosť porozumieť synonymám a povýšilo tak analýzu kľúčových slov na novú úroveň. (Google Algorithm Change History 205612) Zároveň kladlo väčší dôraz na slová ako je dôvera, autorita a reputácia. Aktualizácia ukázala, že kľúčom je poskytovanie relevantných informácií. Väčšia vážnosť bola pripísaná kvalite obsahu webovej stránky. (Google Updates Log 2011)

Obrovskou udalosťou v histórii Google bol v auguste predaj 19M akcií, zvýšenie kapitálu na \$1,67B a nastavenie trhovej hodnoty na viac ako \$20B. Do januára 2005 sa ceny jeho akcií viac než zdvojnásobili. (Google Algorithm Change History 2012)

## **5.4 Aktualizácie roku 2005**

### **5.4.1 NoFollow**

V tom istom čase predstavil Google spolu so spoločnosťami Microsoft a Yahoo! atribút tzv. Nofollow pre boj proti spamu a kontrolu kvality odchádzajúcich odkazov. Tento atribút pomohol odstrániť neoverené odkazy vrátane spamových komentárov bloggov. Aj keď tento zásah nebol priamou aktualizáciou algoritmu, získal postupne výrazný vplyv na graf odkazov.

### **5.4.2 Allegra**

Vo februári 2005 síce správcovia webových stránok zaznamenali v rankingu zmeny, no špecifiká tejto aktualizácie boli nejasné. Niektorí sa domnievali, že na to mala vplyv aktualizácia Allegra, iní to pokladali za dôsledok vylepšenia LSI. Vznikli aj podozrenia, že Google začal podozrivé odkazy penalizovať. (Google Algorithm Change History 2012)

Aktualizácia Allegra pripomínala aktualizáciu Florida, pretože sa opäť venovala znižovaniu vplyvu neprirodzene získaných odkazov a relevancii meta dát. Zároveň sa od tejto aktualizácie stal populárny termín "pre-optimalizovaný" (over-optimized), ktorý popisoval presycovanie kľúčovými slovami. Táto aktualizácia však nemala takmer žiadny dopad na väčšinu rankingov, pretože už po aktualizácii Florida mnoho SEO expertov prestalo používať tieto nepovolené a ľahko odhaliteľné praktiky získavania odkazov. (Google Updates Through The Years 2011)



Allegra bola zároveň pokusom Googlu o identifikáciu spamových stránok, ktoré stále dokázali docieľiť, aby mali vo výsledkoch vyhľadávania vysoký ranking. Google požiadal používateľov o spätnú väzbu ku webovým sídlam, ktoré si naozaj zaslúžili vyšší ranking, ale nedostali ho. Používatelia sa začali sťažovať, že ich stránky zmizli z výsledkov vyhľadávania, hoci niektoré spamové weby majú stále vysoké ranky. (Google Updates Log 2011)

### **5.4.3 Bourbon**

V máji 2005 niekto zo spoločnosti Google (pravdepodobne Matt Cutts) oznámil, určité zmeny v kvalite vyhľadávania 3.5. Nebolo isté, aký druh zmien to bol, ale správcovia webov sa domnievali, že aktualizácia Bourbon zmenila spôsob zaobchádzania s duplicitným obsahom a URL adresami začínajúcimi „www“ a adresami bez „www“ (non-canonical URLs). (Hock c2007)

Pomenovaná bola podľa ulice Bourbon v New Orleans , kde sa v roku 2005 konala Svetová konferencia webmasterov. Bol to vlastne prvý pokus Googlu o odfiltrovanie obsahu, ktorého text bol ukradnutý z inej stránky za účelom rýchleho vytvorenia webu. Tento obsah začal byť následne považovaný za spam. Bol to síce škodlivý zásah, no stránky s ukradnutým obsahom predstavujú problém do dnes. (Google Updates Through The Years 2011)

Bourbon bola zároveň odpoveďou na sťažnosti na spam a požiadavky na opätovné začlenenie stránok do výsledkov vyhľadávania. Pre zvýšenie efektivity boli do procesu zavedené isté strategické zmeny. Aktualizácia sa zameriavala aj na premiestňovanie dát zo starých datacenter do nových. (Google Updates Log 2011)

Do centra pozornosti sa dostali tzv. rodiny webových sídel (web site families, link wheels) a začali byť hlásené prípady neprirodzeného získavania odkazov, keď napríklad inak statická stránka odrazu získala stovky odkazov. (Google Updates Through The Years 2011)

### **5.4.4 Sitemaps XML**

V polovici roku 2005 Google umožnil správcovi webových sídiel vkladať ich mapy sídiel vo formáte XML (XML Sitemaps) prostredníctvom nástrojov správcu webu (Webmasters Tools), čím obišli tradičné mapy v HTML, a získali tak priamy vplyv (aj keď len minimálny) na prechádzanie a indexovanie svojich webov. (Hock c2007)

#### **5.4.5 Personalized Search**

Na rozdiel od predchádzajúcich pokusov o personalizáciu, ktorá vyžadovala vlastné nastavenia a profily, sa v júni 2005 objavila už druhá aktualizácia v tomto mesiaci – personalizované vyhľadávanie (Personalized Search) zamerané priamo na históriu vyhľadávania užívateľov a následného automatického prispôsobenia vyhľadaných výsledkov. Aj keď bol dopad spočiatku veľmi malý, Google naďalej využíval históriu vyhľadávania pre množstvo aplikácií. (Gupta 2012)

#### **5.4.6 Gilligan**

V septembri toho istého roku weboví správcovia spozorovali zopár ďalších zmien (tzv. zásah Gilligan), no spoločnosť Google tvrdila, že ku žiadnej aktualizácii základného algoritmu nedošlo. Matt Cutts v príspevku na blogu situáciu objasnil tým, že Google v tom čase aktualizoval údaje indexu denne, no lišta nástrojov PageRanku a niekoľko ďalších metrík len raz za tri mesiace. (Hock c2007)

#### **5.4.7 Google Local / Maps**

Google Maps - Po marcovom spustení Lokálneho obchodného centra (Local Business Center, LBC) a nasledujúcemu podnecovaniu podnikov k aktualizáciám svojich údajov, začlenil Google svoje údaje z máp do LBC, čo malo podstatný vplyv na celý svet SEO. (Major Google Algorithm Updates of The Last Decade 2012)

#### **5.4.8 Jagger**

V októbri spustil Google sériu aktualizácií, z veľkej časti zameraných na menej kvalitné odkazy vrátane obojstranných odkazov, odkazových fariem a platených odkazov. Aktualizácia Jagger prebiehala najmenej v troch etapách, približne od septembra do novembra 2005, pričom najväčší vplyv dosiahla v októbri. (Gupta 2012)

Google vyzval používateľov k spätnej väzbe ohľadom webových sídel, ktoré používali zakázané SEO praktiky za účelom získania vyššieho rankingu. Weby, ktorým bolo dokázané, že takéto techniky používali, boli z výsledkov vyhľadávania odstránené. Google odstránil kanonické problémy a zvýšenú pozornosť venoval relevancii vzájomných odkazov. (Google Updates Log 2011)

Jagger bola v oblasti SEO významnou aktualizáciou a spustila množstvo ďalších zmien. Od kedy sa do pozornosti dostali on-page faktory, najdôležitejšou otázkou sa stali

prichádzajúce odkazy. Aktualizácia Jagger vyzdvihla význam relevantných odkazov a znížila hodnotu zdanlivo nesúvisiacich náhodných odkazov. To sa dotýkalo aj otázky vzájomných odkazov, ktorou sa Google zaoberal. Táto devalvácia nerecipročných odkazov spustila lavínu ďalších zmien a negatívne postoje k spoločnosti Google, ale keď sa situácia trochu ustálila, bola aktualizácia uznaná ako významné zlepšenie vyhľadávaných výsledkov. Zároveň vznikla možnosť nahlásiť nekvalitné webové sídla, na ktorých boli spustené reklamy typu AdSense v rozpore s pravidlami Googlu. (Google Updates Through The Years 2011)

#### **5.4.9 Big Daddy**

Niekoľko mesačná aktualizácia Big Daddy bola, v technickom pojmání, aktualizáciou infraštruktúry, a trvala od decembra 2005 až do marca ďalšieho roku. Zmenila spôsob zaobchádzania s kanonickými URL adresami, presmerovania (301/302) a ďalšie technické záležitosti. (Hock c2007)

Okrem toho sa Google zameral aj na prichádzajúce a odchádzajúce odkazy, a tak stránky, ktoré mali nedôveryhodné odkazy alebo samy odkazovali na viacero spamových stránok, boli z indexu odstránené, a spamové stránky boli premiestnené do doplnkovej kategórie výsledkov vyhľadávania (Google Updates Log 2011)

### **5.5 Aktualizácie roku 2006**

#### **5.5.1 Supplemental**

V novembri roku 2006 boli vykonané zmeny v spôsobe, akým boli stránky filtrované. Tento krok spočiatku pôsobil ako penalizácia, Google ho však za ňu nepovažoval.

#### **5.5.2 False Alarm**

V decembri roku 2006 sa objavili zmienky o obrovských zmenách v hodnotení výsledkov, ktoré však spoločnosť Google poprela, išlo zrejme o falošný poplach. (Gupta 2012)

## **5.6 Aktualizácie roku 2007**

### **5.6.1 Googlebomb**

Pobláznenie Google bomb z roku 2007 ukázalo svetu význam anchor textu, zároveň však bolo toto obdobie poznamenané politickým útokom. Napríklad pri zadaní výrazu „miserable failure“ (biedne zlyhanie) do okna vyhľadávania sa na prvej pozícii objavila stránka whitehouse.gov, čo bolo spôsobené desiatkami tisícov webmastrov, ktorí nesympatizovali s Georgom Bushom a jeho ideami a preto odkazovali na web Bieleho domu týmto anchor textom. Ďalšia Google bomba mala za následok zobrazovanie stránky Johna Kerryho na prvej pozícii pri hľadaní slova waffle. Po zverejnení týchto udalostí na politickom blogu Google zakročil zmazaním týchto odkazov. (Google Updates Through The Years 2011)

### **5.6.2 Universal Search (aka Google 2.0)**

Aktualizácia Universal Search v máji roku 2007 priniesla koncept vertikálneho vyhľadávania, ktoré je schopné vyhľadať všetky druhy obsahu, ako sú blogy, obrázky, články, správy alebo videá na akúkoľvek konkrétne vymedzenú tému. Išlo v podstate o prínos pre koncových užívateľov bez masívneho zásahu vlastníkov webových sídel. (Beswick 2012)

Tento zásah síce pozmenil formát vyhľadávača Google, no nešlo o typickú zmenu algoritmu.

### **5.6.3 Buffy**

Aktualizácia Buffy bola pomenovaná na počesť Vanessy Fox, ktorá opustila spoločnosť Google. Konkrétnymi zmenami si však nikto nebol úplne istý a Matt Cutts označil túto aktualizáciu len za skupinu menších zmien. (Google Algorithm Change History 2012)

## **5.7 Aktualizácie roku 2008**

### **5.7.1 Dewey**

Hoci sa na prelome marca a apríla 2008 zdalo, že v algoritme Google nastávajú ďalšie významné zmeny, konkrétna povaha aktualizácie Dewey nebola nikdy objasnená. Vznikli však domnienky, že Google pracoval na vlastnom vylepšení služieb, napríklad Google Books.

### **5.7.2 Google Suggest**

V auguste roku 2008 boli predstavené návrhy vyhľadávaní, ktoré sa zobrazujú v rolovacom zozname priamo pod oknom na zadanie hľadanej požiadavky, ihneď ako používateľ zadáva do okna hľadané heslo.

## **5.8 Aktualizácie roku 2009**

### **5.8.1 Vince**

Ďalšou významnou zmenou bola vo februári 2009 aktualizácia Vince, ktorá mala podľa expertov na optimalizáciu stránok výrazne pozitívny vplyv na veľké obchodné značky. (Gupta 2012) Jej cieľom bolo potrestanie stránok, ktoré využívali známe identity veľkých spoločností a odstránenie nekvalitných stránok s chudobným obsahom z prvej stránky výsledkov hľadania. (Beswick 2012)

Hoci Matt Cutts informoval, že išlo len o nepatrné zmeny (Gupta 2012), táto aktualizácia zlepšila kvalitu vyhľadávania pre koncových užívateľov a viedla k modernému vnímaniu autorít. (Beswick 2012)

### **5.8.2 Rel-Canonical Tag**

Približne v tom istom čase vyhlásila spoločnosť Google spolu so spoločnosťami Microsoft a Yahoo! podporu tzv. kanonického tagu (Canonical Tag, <rel="canonical">) (Gupta 2012), ktorý webovým správcom umožnil posielat' robotom kanonizačné signály bez zásahu návštevníkov. (Hock c2007)

### **5.8.3 Caffeine**

Google v auguste 2009 poskytol náhľad obsiahlej zmeny infraštruktúry – Caffeine (Preview) vytvorenej za účelom vylepšenia a zrýchlenia procesu prechádzania stránok. (Gupta 2012)

To umožnilo hlbšie spracovanie obsahu stránky, a tým aj nájdenie relevantnejších výsledkov k danej požiadavke. Vďaka tejto aktualizácii označil Google rýchlosť načítania stránky za ďalší z rankingových faktorov. (Google Updates Log 2011)

Okrem toho Caffeine priniesol aj rozšírenie indexu a s cieľom indexovania a hodnotenia stránok takmer v reálnom čase. Časový plán bol rozdelený po mesiacoch s konečným spustením v Spojených štátoch začiatkom roku 2010 a predpokladaným trvaním do leta (Hock c2007), jeho spustenie ale bolo niekoľkokrát odložené, a nepodarilo sa ho ani spustiť podľa plánu celosvetovo naraz. (Google Updates Log 2011)

#### **5.8.4 Real-Time Search**

Koncom roku 2009 bolo umožnené vyhľadávanie v reálnom čase. Príspevky na Twitteri, Google správy, naposledy indexovaný obsah a niekoľko ďalších zdrojov slúžilo pre zobrazovanie príspevkov na niektorých stránkach s výsledkami vyhľadávania v reálnom čase. Postupne sa tieto zdroje rozrastali a zahrnuli aj sociálne médiá. (Hock c2007)

### **5.9 Aktualizácie roku 2010**

#### **5.9.1 Google Places**

Google miesta (Google Places) boli zo začiatku, od roku 2009, len súčasťou Google máp a oficiálne boli spustené v apríli 2010. To prinieslo nové lokálne výsledky vyhľadávania a vznikol priestor pre nové lokálne možnosti reklamy. (Gupta 2012)

Rok 2009 zaznamenal veľký počet zmien v algoritme, podľa Matta Cuttsa ich bolo asi 300 až 550. (Google Dance Infographic finished 2011)

#### **5.9.2 May Day**

Aktualizácia May Day sa objavila na prelome apríla a mája roku 2010 (Gupta 2012) a získala si označenie „deň, kedy zomrel long tail (viacslovné slovné spojenia hľadané vo vyhľadávačoch)“. Mnoho webových sídiel prišlo o 5-15% bežnej návštevnosti z týchto viacslovných dotazov. (Google Dance Infographic finished 2011) Prepojenie stránok odkazmi, bez ohľadu na hodnotu odkazov, sa tak stalo jedným z najdôležitejších faktorov vplyvajúcich na poradie stránok vo výsledkoch vyhľadávania. (Google Updates Log 2011)

V tomto období správcovia webov zaznamenali výrazné poklesy v dlhodobej prevádzke. Matt Cutts neskôr potvrdil, že zmena algoritmu – May Day spôsobila tento dopad z dlhodobého hľadiska. (Hock c2007)

### **5.9.3 Caffeine (Rollout)**

Po dlhých mesiacov testovania Google spustil v júni 2010 aktualizáciu Caffeine (Gupta 2012), ktorá sľubovala až o 50% čerstvejšie výsledky vyhľadávania. Zabrala takmer 100 miliónov gigabytov úložného miesta a každým dňom tento objem narastal minimálne o 100 000 GB. (Google Dance Infographic finished 2011) Zvýšila sa rýchlosť vyhľadávania a bol upravený proces prechádzania a indexovania stránok. (Gupta 2012)

Pri používaní vyhľadávača Google nie sú výsledky hľadané cez www (World Wide Web), ale vyhľadávajú sa v knižnici stránok indexu Google. Aktualizácia Caffeine predstavovala zmenu v jeho infraštruktúre, na základe ktorej sa urýchléné indexovanie nových a updatovaných stránok prejavilo v poskytovaní podstatne aktuálnejších výsledkov pri vyhľadávaní. (Google Updates Through The Years 2011)

### **5.9.4 Brand update**

V auguste 2010 povolil Google viacnásobné zobrazenie jednej domény vo výsledkoch vyhľadávania (pôvodne sa každá doména mohla na stránke s výsledkami vyhľadávania zobrazit' len jeden až dvakrát). (Hock c2007)

Táto zmena však nepredstavuje bežnú aktualizáciu algoritmu.

### **5.9.5 Google Instant**

V septembri bola zadaná požiadavka na rozšírenie našeptávača (Google Suggest) označovaná ako Google Instant, ktorá spočívala v zobrazovaní výsledkov hľadaného hesla už v priebehu jeho zadávania používateľom do okna vyhľadávača. (Gupta 2012)

Okrem toho, že bol tento vyhľadávací nástroj prínosný pre tých, ktorí presne nevedeli, čo hľadajú (používatelia tak mohli ušetriť až 2 sekundy pri každom vyhľadávaní), priniesla táto zmena aj niekoľko malých zmien v oblasti optimalizácie webových stránok. Marketingoví pracovníci sa spočiatku obávali, že budú zasiahnuté kľúčové frázy skladajúce sa z niekoľkých slov (long tails), pretože používatelia ich neuvidia medzi prvými návrhmi. Aj keď to bola to istej miery pravda, skutočnosť, že Instant ukladal časy hľadania, sa v mnohých prípadoch prejavila v relevantnejších výsledkoch, a nemala dopad na špecializované stránky a menšie weby, ktoré aj naďalej uprednostňovali kvalitu. (Beswick 2012)

### **5.9.6 Instant Previews**

Koncom roka, v novembri 2010, sa vo výsledkoch hľadania objavila ikona lupy, ktorá užívateľom umožňovala rýchly náhľad vstupných stránok priamo na stránke s výsledkami bez toho, aby nájdené odkazy museli byť otvorené. (Gupta 2012)

To signalizovalo opätovné zameranie sa Googlu na kvalitu, dizajn a použiteľnosť vstupných stránok.

### **5.9.7 Social Signals**

V decembri Google a Bing potvrdili, že pri výpočte rankingu využívajú aj sociálne prvky vrátane údajov z Twitteru a Facebooku. Podľa Matta Cuttsa to bola súčasť blížiaceho sa vývoja Google praktík.

### **5.9.8 Negative Reviews**

Na prípad uverejnený v New York Times, v ktorom ranking stránky elektronického obchodu DecorMyEyes negatívne ovplyvnili recenzie, Google reagoval ďalšou úpravou algoritmu. (Hock c2007)

Táto úprava, nazvaná Décor, znížila v decembri 2010 hodnotenie stránok, ktoré sa zámerne snažili vyprovokovať návštevníkov k tomu, aby blogovali o svojom negatívnom postoji a znechutení, pričom poskytli odkaz na danú stránku. Predtým neboli z kvalitných stránok žiadne relevantné negatívne odkazy, bez ohľadu na dôvod odkazu. (Googlebot 2012)

## ***5.10 Aktualizácie roku 2011***

### **5.10.1 Penalizácia Overstock.com**

V januári 2011 Google verejne penalizoval spoločnosť Overstock.com za využívanie pochybných SEO praktík za účelom výrazného zvýšenia rankingov. (Gupta 2012)

Spoločnosť JCPenney bola za podobné správanie pokutovaná o mesiac neskôr. Obe tieto situácie sa prejavili posunom v postoji Google a stali sa predzvest'ou novej veľkej aktualizácie. (Hock c2007)



### **5.10.2 Attribution Update**

Ako odpoveď na početné prípady, spoločne označované ako „vysoko profilový spam“, spustila spoločnosť Google ďalšiu aktualizáciu. Označovaná bola ako atribučná aktualizácia (Gupta 2012) a bola zameraná na lepšie hodnotenie obsahu, aby sa zamedzilo jeho kradnutiu z cudzích webových sídiel. Podľa Matta Cuttsa ovplyvnila približne 2% vyhľadávaní a stala sa jasným predchodcom aktualizácií Panda. (Hock c2007)

### **5.10.3 Update Panda (známa aj ako aktualizácia Farmer)**

Táto významná aktualizácia algoritmu sa hlboko dotkla webových sídiel a ovplyvnila až 12% výsledkov vyhľadávania (toto číslo pochádza priamo od spoločnosti Google). (Kluver 2012)

Zámerom tejto aktualizácie bolo vysporiadanie sa s mnohými otázkami týkajúcimi sa kvality webových sídiel, ako sú obsahové farmy, chudobný obsah, vysoký pomer reklám vzhľadom k obsahu a ďalšie. (Gupta 2012) Hoci bola Panda spustená už 23. februára 2011, Európu zasiahla až v apríli. (Hock c2007)

### **5.10.4 Tlačidlo +1**

Spoločnosť Google vytvorila v marci 2011 tlačidlo +1, ktoré umiestnila hneď vedľa odkazov vo výsledkoch vyhľadávania. Reagovala tak zároveň na svojich konkurentov – napríklad Facebook a Twitter, ktorý podobný štýl hodnotenia už používali. (Gupta 2012)

Kliknutím na ikonu "+1" tak mohli užívatelia ovplyvniť organické aj platené výsledky vyhľadávania vo svojom sociálnom prostredí. (Hock c2007)

### **5.10.5 Panda 2.0**

V apríli 2011 Google rozšíril aktualizáciu Panda na všetky anglické požiadavky celosvetovo - bez obmedzenia na anglicky hovoriace krajiny. Zahrnuté boli aj nové faktory, ako napríklad údaje o stránkach blokových užívateľmi prostredníctvom prehliadača Chrome alebo priamo na stránke s výsledkami hľadania. (Kluver 2012)

### **5.10.6 Panda 2.1**

Ďalšia séria zmien (spočiatku označovaná ako Panda 3.0) (Hock c2007)

vykonaných v rámci aktualizácie Panda, ktorej detaily však nikdy neboli zo strany Google zverejnené a prejavila sa len v minimálnom rozsahu.

#### **5.10.7 Schema.org**

Google spolu so spoločnosťami Microsoft a Yahoo! spojili svoje sily za účelom zjednodušenia práce webmasterov tým, že (Gupta 2012) v júni roku 2011 vyhlásili podporu konsolidovaného prístupu k štruktúrovaným údajom. (Hock c2007)

Vytvorili teda súbor schém – html tagov, ktoré mali okrem iného zaistiť aj obsiahlejšie a kvalitnejšie výsledky hľadania. (Gupta 2012)

#### **5.10.8 Panda 2.2**

21. júna Google pokračoval v aktualizácii stránok a údajov, ktoré Panda ovplyvnila, tentokrát však zmeny neprebiehali v reálnom čase a boli vykonané oddelene od hlavného indexu, čím pripomínali prvé aktualizácie Google Dance. (Hock c2007)

#### **5.10.9 Google+**

Po predchádzajúcich neúspešných pokusoch spustil Google koncom júna svoju vlastnú sociálnu sieť, Google Plus, ktorá sa stala vážnou hrozbou pre konkurenčný Facebook. (Gupta 2012)

Google+ vynikal tzv. kruhmi (Circles), ktoré slúžili na zdieľanie obsahu a boli úzko prepojené s inými produktmi, ako napríklad Gmail. V priebehu prvých dvoch týždňov bolo do siete prihlásených 10 miliónov používateľov.

#### **5.10.10 Panda 2.3**

V júli 2011 sa weboví správcovia domnievali, že Google spustil ďalšiu aktualizáciu, no nebolo isté, či sa prejavili nejaké nové faktory, alebo išlo len o aktualizáciu údajov a hodnotiacich faktorov updatu Panda.

#### **5.10.11 Panda 2.4**

Spoločnosť Google spustila 12. Augusta 2011 update Panda na medzinárodnej úrovni – v rámci globálnych dotazov v anglickom jazyku a pre požiadavky v iných jazykoch okrem čínskeho, japonského a kórejského jazyka. (Hock c2007)

Táto zmena mala pravdepodobne dopad na 6-9 % vyhľadávacích dotazov v krajinách, ktoré boli zasiahnuté. (Gupta 2012)

#### **5.10.12 Rozšírenie vnorených odkazov (Sitelinks)**

Po krátkom experimentovaní Google oficiálne spustil rozšírenie vnorených odkazov predovšetkým pre vyhľadávania značiek. Spočiatku išlo o zoskupenia po 12 odkazoch, no krátko po spustení Google obmedzil ich počet na 6 odkazov.

#### **5.10.13 Prvky číslovania stránok**

Aby sa predišlo problémom prezerania stránok a duplikáciám spôsobeným číslovaním stránok, predstavil Google atribúty odkazov rel=“next“ a rel=“prev“. Taktiež oznámil vylepšenie automatickej kanonizácie a konsolidácie stránok pri voľbe „Zobraziť všetky“ („View All“).

#### **5.10.14 516 aktualizácií**

Eric Schmidt, generálny riaditeľ spoločnosti Google, dňa 21. Septembra roku 2011 zverejní, že v roku 2010 bolo vykonaných 516 aktualizácií, hoci testovaných bolo viac než 13 tisíc. (Hock c2007)

#### **5.10.15 Panda 2.5**

Koncom mesiaca bol spustený ďalší update Panda s označením Panda 2.5. Špecifické detaily toho, čo sa vlastne zmenilo neboli jasné, no niektoré webové sídla ohlásili obrovské straty. (Gupta 2012)

#### **5.10.16 Panda „Flux“**

Dňa 5. októbra Matt Cutts na Twitteri oznámil, že v priebehu niekoľkých týždňov je možné očakávať zmeny spojené s aktualizáciou Panda. Ďalšie malé Panda updaty sa objavili 3. a 13. októbra a 18. novembra. (Hock c2007)

#### **5.10.17 Šifrovanie požiadaviek**

Ešte v tom istom mesiaci Google ohlásil, že vyhľadávacie požiadavky budú z dôvodu súkromia šifrované, čo negatívne ovplyvnilo dáta o príchodoch prostredníctvom kľúčových slov. (Gupta 2012)

#### **5.10.18 Aktualizácia čerstvosti (Freshness update)**

Začiatkom novembra Google informoval o ďalšej úprave algoritmu (Gupta 2012), na ktorú SEO experti čakali už dlhšiu dobu. (Hock c2007) Táto úprava pozitívne odmenila aktuálny a čerstvý obsah a zasiahla až 25% výsledkov.

#### **5.10.19 Balíček 10 aktualizácií**

So zámerom zvýšenia transparentnosti zverejnil Matt Cutts zoznam posledných desiatich zmien algoritmu. Mnohé z nich však boli príliš nevýrazné a chýbalo im konkrétne časové určenie.

#### **5.10.20 Panda 3.1**

Po update Panda 2.5 nastalo obdobie vplyvu Panda Flux, v ktorom prebehlo niekoľko aktualizácií algoritmu bez výraznejších zmien. Hoci Panda 3.0 nebola nikdy oficiálne vyhlásená, tieto zmeny analytici nazývali Panda 3.1.

#### **5.10.21 Decembrový balíček 10 aktualizácií**

Prvý decembrový deň roku 2011 vyšiel zoznam ďalších 10 updatov bez časovej determinácie, medzi ktorými bolo napríklad odhaľovanie zakúpených domén, ktoré neobsahovali žiaden obsah (parked domains), overovanie čerstvosti vyhládaných blogov a obrázkov, a niekoľko ďalších. (Gupta 2012)

### ***5.11 Aktualizácie roku 2012***

#### **5.11.1 Januárový balíček 30 aktualizácií**

Zoznam z 5. januára 2012 obsahoval 30 aktualizácií za predchádzajúci mesiac, okrem iného sa v ňom objavili aj nové rozširujúce informácie používané vo výsledkoch vyhľadávania (Rich Snippets) a vylepšenia v oblasti odkazov na relevantné stránky. (Gupta 2012) Hranice medzi aktualizáciami algoritmu a novými vlastnosťami systému boli čoraz menej jednoznačné. (Google Algorithm Change History 2012)

#### **5.11.2 Search + Your World**

Obrovské zmeny v oblasti personalizácie vyhlásila spoločnosť Google 10. januára 2012. Ich podstata spočívala v zahrnutí užívateľských profilov a ďalších sociálnych dát do stránky obsahujúcej vyhľadane výsledky. Stránka obsahovala aj tlačidlo, ktoré umožňovalo túto personalizáciu vypnúť. (Gupta 2012)

### **5.11.3 Panda 3.2**

Aktualizácia Panda patrí bezpochyby k updatom, ktoré najviac zatriasli výsledkami vyhľadávania. Hoci spoločnosť Google jasne nevysvetlila, čo konkrétne sa v rámci Pandy zmenilo, potvrdila ale, že zmeny v oblasti Pandy boli vykonané aj 18. januára 2012. (Dholakiya 2013)

### **5.11.4 Reklamné odkazy nad prehybom stránky (Ads above the Fold)**

Začiatkom roku 2012 bola na základe novej aktualizácie znížená hodnota stránok obsahujúcich príliš veľký počet reklám na stránke na úkor obsahu, čím sa potvrdilo, že sa usporiadanie prvkov na stránke stáva neustále dôležitejším kritériom. Tento krok však nedostal žiadny oficiálny názov a špekulovalo sa aj o tom, že podobná akcia bola už vykonaná v rámci update Panda. (Gupta 2012)

### **5.11.5 Venice**

Presný dátum ďalšej aktualizácie v poradí nebol nikdy určený, Google jeho označenie „Venice“ spomenul len ako súčasť ďalšej mesačnej aktualizácie. (Google Algorithm Change History 2012) Tento update si nezískal dostatok pozornosti a uznania aj napriek tomu, že v skutočnosti ovplyvnil výsledky týkajúce sa každodenných užívateľov vo väčšej miere ako veľká časť ostatných aktualizácií. Spustený bol vo februári v rámci viac ako 40 updatov obsiahnutých v jednom kompaktnom balení, cieľený bol na úpravu a zdokonalenie lokálnych výsledkov vyhľadávania. Tie sú vo veľkej miere závislé od toho, kde sa používateľ nachádza v národnom alebo regionálnom zmysle. Venice bol do značnej miery zatienený tým, čím sa následne stala aktualizácia Penguin. Bez ohľadu na to však výrazne pomohol kvalite vyhľadávania a použiteľnosti pre bežných užívateľov. (Beswick 2012)

### **5.11.6 Balíček 40 aktualizácií**

Zoznam ďalších výrazných zmien zvýšenia kvality vyhľadávania Google zverejnil 27. februára 2012. Obsahoval údajne okolo 40 aktualizácií, ktoré sa objavili v priebehu tohto mesiaca, z ktorých boli najvýznamnejšie mnohopočetné zmeny týkajúce sa vyhľadávania

v obrázkoch, aktualizácie čerstvosti spolu s vyradením dvoch starých častí algoritmu a samozrejme aktualizácia Panda. (Google Algorithm Change History 2012)

#### **5.11.7 Panda 3.3**

Update Panda sa stal najdlhšie pretrvávajúcim updatom – Panda s označením 3.3 bola spustená presne tri dni po výročí prvej aktualizácie Panda. (Dholakiya 2013)

#### **5.11.8 Video o kvalite vyhľadávania**

V polovici marca síce nešlo o ďalšiu aktualizáciu algoritmu, ale spoločnosť Google zverejnila vzácny príspevok zo stretnutia zameraného na proces vyhľadávania a jeho kvalitu, čo poskytlo jedinečnú príležitosť pre nazretie do procesov a priorít vyhľadávacej technológie Google.

#### **5.11.9 Panda 3.4**

Ďalšia séria updatov Panda bola vyhlásená na Twitteri 23. marca toho istého roku a odhadovalo sa, že ovplyvnila okolo 1,6% výsledkov na zobrazených na výsledných stránkach vyhľadávačov.

#### **5.11.10 Trest za prehnané optimalizovanie stránky**

Každá z aktualizácií Googlu je pomenovaná ľuďmi z odvetvia a jedna z marcových bola nazvaná Aktualizácia trestu za pre-optimalizovanie stránky (Over Optimization Penalty). Bližšie špecifiká nie sú známe, no v rámci SEO sa týkala anchor textu. Každý, kto odkazuje na web, ktorý obsahuje niekoľkokrát použitý rovnaký anchor text, v snahe urobiť určitú kľúčovú frázu cieľovou, dosiahne veľmi slabé výsledky. (Hock c2007)

#### **5.11.11 Marcový balíček 50 aktualizácií**

Začiatkom apríla sa objavil zoznam ďalších významných aktualizácií obsahujúci 50 aktualizácií pripadajúcich na mesiac marec 2012. V rámci neho bola potvrdená aktualizácia Panda 3.4, zmeny v ohodnocovaní anchor textu, update obrázkového vyhľadávania a spôsob, akým sú interpretované lokálne zamerané hľadané výrazy. (Google Algorithm Change History 2012)

#### **5.11.12 Chyba v hodnotení parkovaných domén (Parked Domain Bug)**

Chyba nesprávne zamenená za úmyselnú zmenu algoritmu spôsobila rozruch medzi správcami webových sídiel, ktorí reportovali drastické zmeny v rankingoch. Google potvrdil, že omyl v dátach spôsobil, že niektoré weby boli neprávom považované za parkované domény a preto ich hodnotenie rapídne kleslo. Nešlo však o úmyselnú zmenu algoritmu. (Dholakiya 2013)

#### **5.11.13 Panda 3.5**

O pár dní neskôr spustil Google ďalšiu aktualizáciu Panda, ktorá spočívala v niekoľkých zmenách, a preto bolo zložité presne zmerať jej dopad. Pravdepodobne však bola len ďalším rutinným updatom s minimálnym vplyvom.

#### **5.11.14 Penguin**

Niekoľko týždňov po špekuláciách ohľadom aktualizácie trestov za prehnanú optimalizáciu spustil Google tzv. Webspam Update, ktorý bol o krátky čas premenovaný na Penguin.

Hoci tento problém pretrvával a postupne sa vyvíjal už od počiatku webového vyhľadávania, snaha o ovplyvňovanie systému PageRank dosiahla svoj vrchol približne v tomto čase. Odpoveďou bolo práve spustenie updatu Penguin v apríli 2012, ktorý opäť kládol dôraz na autoritu, zastavil vývoj schém miliónov odkazov, záludných maskovacích trikov a praktík zameraných na kľúčové slová (napr. v anchor textoch), ktoré škodili poctivým správcam webov. (Beswick 2012) Najvýraznejšie sa teda Penguin prejavil v úprave nastavení niekoľkých spamových faktorov, napríklad preplňovania kľúčovými slovami (keyword stuffing), a zasiahol približne 3,1% vyhľadávacích dotazov v angličtine. (Google Algorithm Change History 2012)

#### **5.11.15 Panda 3.6**

Necelý týždeň po spustení aktualizácie Panda 3.5 bola spustená Panda 3.6. Jej následky neboli príliš zreteľné a predpokladá sa len minimálny vplyv na hodnotenie výsledkov. (Dholakiya 2013)

#### **5.11.16 Aprílový balíček aktualizácií**

Balíček updatov za mesiac apríl obsahoval 52 aktualizácií vykonaných v priebehu tohto mesiaca. Zverejnený bol 4. mája a boli v ňom obsiahnuté okrem iných aj zmeny súvisiace s aktualizáciou Penguin. Medzi jeho ďalšie významné prvky patrilo zväčšenie základne



indexu o 15%, vylepšenie číslovania stránok a niekoľko aktualizácií odkazov na podstránky (sitelinks). (Google Algorithm Change History 2012)

#### **5.11.17 Knowledge Graph**

V polovici mája toho istého roku sa urobil Google ďalší krok smerom k sémantickému vyhľadávaniu predstavením funkcie Knowledge Graph. Táto funkcia v sebe zahŕňala pridané objekty zobrazené v pravej časti stránky s výsledkami hľadania a ukazovala doplnkové informácie o určitých témach ako sú miesta, ľudia alebo veci. Zároveň sa plánovalo postupné zobrazovanie týchto panelov na čoraz väčšom počte stránok. (Dholakiya 2013)

Tento spôsob vyhľadávania zmenil pohľad na klasické vyhľadávanie používajúce obrovskú množinu kľúčových slov, a ukázal súvislosti medzi nimi a ľuďmi, zvieratami, budovami, krajinami a pod., o ktorých pozná viac než tri a pol miliardy základných údajov. Tým sa medzi objektmi vytvorili vzájomné vzťahy, čo pomáhalo aj pri rozlišovaní medzi objektmi s rovnakým pomenovaním. (Janouš 2012)

#### **5.11.18 Penguin 1.1**

25. mája spustil Google po aktualizácii algoritmu Penguin prvý update zameraný na dáta. To potvrdilo, že dáta aktualizácie Penguin boli spracovávané mimo hlavného indexu, podobne ako to bolo pri aktualizácii Panda.

#### **5.11.19 Májový balíček aktualizácií**

Balíček najvýznamnejších aktualizácií za mesiac máj obsahoval zoznam 39 aktualizácií, ktoré sa udiali v tomto mesiaci. Medzi najdôležitejšie patril update Penguin, zmeny v Google správach (Google News), lepšia detekcia schém odkazov a úpravy prepisovania názvov a rozširujúcich informácií. (Google Algorithm Change History 2012)

#### **5.11.20 Panda 3.7**

Spustenie aktualizácie Panda 3.7 ovplyvnilo ešte viac výsledkov ako jej predchodcovia - podľa vyjadrenia Google mala vplyv až na 1% hľadaných hesiel, čo je oveľa viac ako to bolo pri jej predchádzajúcich dvoch verziách.

#### **5.11.21 Panda 3.8**

Panda 3.8 ovplyvnila menšie percento výsledkov ako Panda 3.7. a mala povahu výlučne dátového updatu bez akýchkoľvek zmien algoritmu.

#### **5.11.22 Upozornenia na odkazy (Links Warnings)**

Rovnako ako v marci a apríli, dostalo aj 19. júna mnoho webmasterov z Google Webmaster Tools (Google nástroje pre správcov webov) varovanie o neprimerených odkazoch. Spoločnosť však uviedla, že tieto varovania sa netýkali vážnych problémov. (Dholakiya 2013)

#### **5.11.23 Panda 3.9**

Ďalší update Panda sa objavil 24. júla 2012. Správcovia webov spozorovali 5-6 dní trvajúce výkyvy v rankingoch svojich stránok bez výraznejších nárastov alebo poklesov a spoločnosť Google nakoniec oznámila, že update mohol ovplyvniť približne 1% vyhľadávaní.

#### **5.11.24 Balíček 86 aktualizácií za obdobie jún/júl**

Po letnej prestávke bol spoločný zoznam aktualizácií za mesiace jún a júl vydaný v prvej polovici augusta. Medzi najhlavnejšie aktualizácie patrili obnovenia v rámci dát a algoritmu Panda, zmeny zoskupovania webov, výrazný nárast rankingov dôveryhodných zdrojov a pravdepodobne aj vylepšenie funkcie usporadúvania rankov. (Google Algorithm Change History 2012)

#### **5.11.25 Penalizácia DMCA**

Google v auguste oznámil, že plánuje penalizáciu webov, ktoré opakovane porušujú copyright prostredníctvom žiadostí o zrušenie DMCA. Tento update bol plánovaný na deň 13.8.2013 (Dholakiya 2013) a bol nazývaný Penalizácia DMCA alebo Emanuel na počesť talentového agenta Ari Emanuela. Odstraňoval autorsky chránený materiál z výsledkov vyhľadávania na požiadanie. Ak neoriginálny obsah nie je bezpečne priradený jeho tvorcom, bráni tak vytvoreniu čo najlepšieho možného online obsahu, a to je negatívne pre nielen Google, ale aj pre jeho tvorcov a užívateľov. (Beswick 2012)

#### **5.11.26 Stránka vyhľadávača so 7 výsledkami**

14. augusta nastala výrazná zmena na stránke s výsledkami hľadania – pri veľkom počte hľadání sa namiesto 10 prvých výsledkov zobrazilo na stránke len 7 výsledkov. Táto aktualizácia spustená niekoľko dní ovplyvnila okolo 18 percent sledovaných kľúčových slov. (Dholakiya 2013)

#### **5.11.27 Panda 3.9.1**

V druhej polovici augusta spustil Google ešte jednu aktualizáciu dát Panda, ale jej dopad bol relatívne malý. Keď sa sérií updatov Panda 3.0 – 3.9 minuli čísla, tento update bol nazvaný Panda 3.9.1.

#### **5.11.28 Panda 3.9.2**

O mesiac neskôr bola spustená ďalšia aktualizácia Panda, označená 3.9.2, ktorej povaha bola čisto dátová. U rankingov síce nastal určitý posun, no v neporovnateľnej miere oproti minulým updatom.

#### **5.11.29 Panda #20**

Aktualizácia Panda v poradí 20 sa časovo kryla s aktualizáciou Exact-Match Domain. Bola pomerne výrazná (zmeny nastali aj v algoritme aj v dátach) a oficiálne mala vplyv až na 2,4% vyhľadávacích požiadaviek. Od tohto updatu sa už pokračovali v označovaní podľa poradového čísla.

#### **5.11.30 Aktualizácia domén s presnou zhodou (EMD)**

Google oznámil zmenu v zaobchádzaní s doménami s presnou zhodou, čo viedlo ku obrovskému zníženiu hodnotení a počet EMD sa v dátach MozCastu znížil o viac ako 10%. Oficiálne bolo zverejnené, že táto zmena ovplyvnila 0,6% objemu vyhľadávania.

#### **5.11.31 Balíček updatov august/september**

Zoznam aktualizácií za letné mesiace júl a august obsahoval 65 updatov vrátane zmeny zobrazenia 7 výsledkov na stránke vyhľadávača, rozšírenie Knowledge Graph, aktualizácie výpočtu kvality stránky a zmeny v určovaní lokálnych výsledkov. (Google Algorithm Change History 2012)

#### **5.11.32 Penguin 3**

Hoci sa predpokladalo, že ďalšia zo série aktualizácií Penguin bude významná, Google nakoniec spustil malý update Penguin zameraný na dáta, ktorý ovplyvnil 0,3% vyhľadávacích požiadaviek. Penguin začal byť číslovaný, podobne ako Panda, podľa poradia. (Dholakiya 2013)

### **5.11.33 Rozloženie stránky #2**

V októbri 2012 Google vyhlásil aktualizáciu spočívajúcu v zmene na pôvodné rozloženie stránky, ktorá bola zameraná na weby obsahujúce príliš veľa reklám nad prehybom. Nie je síce jasné, či išlo o zmenu algoritmu alebo len o obnovu dát v štýle Pandy.

### **5.11.34 Panda #21**

Päť a pol týždňa po spustení Pandy 20 bola spustená Panda s poradovým číslom 21. Išlo o malý update, ktorý podľa oficiálnych informácií ovplyvnil iba 1,1% požiadaviek v anglickom jazyku.

### **5.11.35 Panda #22**

Po niekoľkých nejasných signáloch Google koncom novembra potvrdil aktualizáciu Panda 22, ktorá bola zameraná výlučne na dáta.

### **5.11.36 Rozšírenie Knowledge Graph**

4. decembra pridal Google funkcionality Knowledge Graph aj pre požiadavky v inom ako anglickom jazyku, vrátane španielskeho, francúzskeho, nemeckého, portugalského, japonského, ruského a talianskeho. Táto aktualizácia bola viac než len prekladom a Knowledge Graphu pridala rozšírené možnosti.

### **5.11.37 Panda #23**

Niekoľko dní pred vianočnými sviatkami bola spustená ďalšia aktualizácia Panda, v poradí už 23. Oficiálne bola označovaná ako „refresh“ a zasiahla 1,3% požiadaviek v anglickom jazyku, čo malo väčší dopad ako predchádzajúce aktualizácie Panda 21 a Panda 22.

## **5.12 Aktualizácie roku 2013**

### **5.12.1 Panda #24**

Prvá oficiálne vyhlásená aktualizácia roku 2013 mala dopad na 1,2% vyhľadávacích požiadaviek.

### **5.12.2 Panda #25**

14. marca 2013 Matt Cutts na Search Marketing Expo avizoval blížiacu sa ďalšiu v poradí aktualizáciu Panda, pričom sa zmienil, že by mohlo ísť o posledný update pred začlenením Pandu do jadra algoritmu. Presný dátum nebol potvrdený, ale predpokladalo sa obdobie okolo 13. a 14. marca. (Google Algorithm Change History 2012)

### **5.13 Budúcnosť webového vyhľadávania**

Vyhľadávač Google, ktorého prioritou je poskytovanie čo najpresnejších a najrelevantnejších výsledkov na hľadaný termín, sa od jeho počiatkov zmenil k nepoznaniu. Vzniklo obrovské množstvo nových služieb, časté zmeny v oblasti algoritmu hodnotenia stránok posúvajú vyhľadávací proces stále bližšie k dokonalosti, kľúčové slová prestávajú byť vnímané len ako termíny, ale vytvárajú sa medzi nimi vzťahy a vzniká obrovská množina štruktúrovaných informácií. K negatívam však patrí napríklad spam a stále sa objavujúce nové pokusy o manipuláciu hodnotenia stránok, s ktorými Google, ale aj ostatné vyhľadávače musia bojovať.

Vývoj Google napredoval takou rýchlosťou, že je len ťažko predstaviteľné, kam sa dostane o niekoľko rokov. Vývoj však nespočíva len vo vylepšovaní vnútorných procesov, ale aj v sledovaní technologického pokroku a reakciou naňho. Napríklad, keď s príchodom mobilných telefónov, ktoré boli schopné pripojenia na internet, vznikol responzívny webdesign a rôzne ďalšie prispôsobenia a funkcie pre využívanie mobilného internetu, Google ihneď reagoval prispôbením niektorých funkcií predovšetkým indexovania pre tieto zariadenia.

Pre potrebu stále sa rozvíjajúcej internetovej komunity, pre ktorú sa kvalitný obsah stáva nenahraditeľným, spustil Google aktualizácie Panda a Penguin. Zároveň sprísnil hodnotenie onpage a offpage faktorov webových stránok a zameral sa aj na sociálne signály, ktoré predtým neboli brané do úvahy s veľkou vážnosťou.

Amid Singhal, senior vice prezident oddelenia vyhľadávania v spoločnosti Google, definoval štyri najzákladnejšie technické výzvy, ktorým Google v súčasnosti čelí. Patrí medzi ne knowledge graph, rozpoznávanie reči, porozumenie prirodzeného jazyka a porozumenie konverzácie. Knowledge Graph je o pochopení vzťahov medzi predmetmi a myšlienkami a uvedomenie si súvislostí medzi nimi a toho, čo majú spoločné. Tieto myšlienky zároveň tvoria hlavnú motiváciu pre vylepšenie vyhľadávania v Google. Pri rozpoznávaní reči je

podstatou premena ľudského hlasu do formy textu, čo by v budúcnosti mohlo slúžiť ako základ hlasového vyhľadávania. Pri problematike prirodzeného jazyka je dôležité porozumenie tých detailov, ktoré umožnia premenu hlasovej konverzie na zmysluplnú informáciu. Konverzácia súvisí s prirodzeným jazykom. (Geron 2013)

Google si svoje plány prísne stráži, no na základe minulého vývoja a náznakov zo strany spoločnosti je takmer isté, že aj rok 2013 bude plný zmien a vylepšení apsoň v takej miere, ako boli roky 2011 a 2012.

Zaujímavým nástrojom predpovedajúcim blížiacu sa zmenu alebo úpravu algoritmu je nástroj Mozcast, ktorý publikuje akúsi „predpoveď počasia“ v rámci Google algoritmu a pomáha tak správcom webových sídel pri posudzovaní zmien v ich rankingoch. Dostupný je z URL adresy <http://mozcast.com/>.

-1998-

SPUSTENIE VYHĽADÁVAČA  
Google

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

-1999-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

-2000-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

GOOGLE TOOLBAR

-2001-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

-2002-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

PRVÁ  
ZDOKUMENTOVANÁ  
AKTUALIZÁCIA

-2003-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

BOSTON

CASSANDRA

DOMINIC

ESMERALDA

FRITZ

DOPLNKOVÝ  
INDEX

FLORIDA

-2004-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.

AUSTIN

BRANDY



-2005-



-2006-



-2007-



-2008-



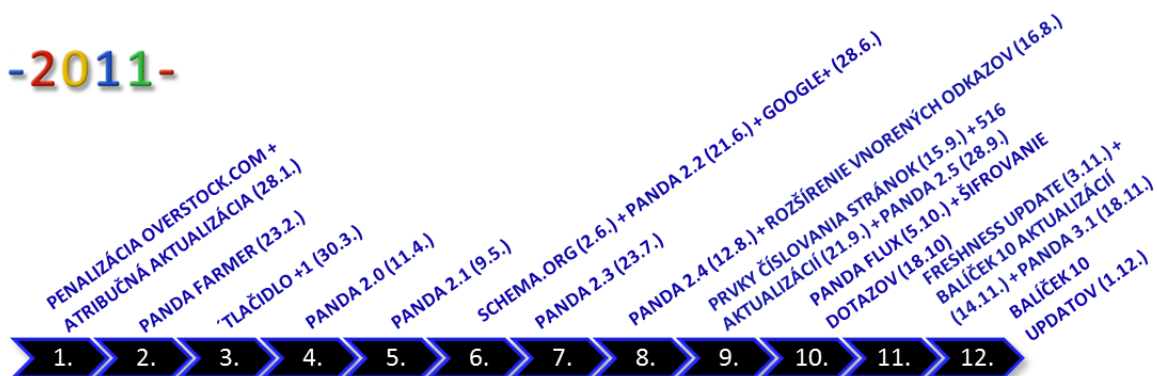
-2009-



-2010-



-2011-



-2012-



-2013-



Obr. č. 5: Časová priamka aktualizácií Google

## ZÁVER

S neustálym nárastom voľne dostupných informácií predovšetkým v oblasti internetu, je nevyhnutné ich spracovanie a usporiadanie. Pôvodné webové vyhľadávače však tieto informácie hodnotili prevažne na základe slov v nich obsiahnutých. Zlom prišiel s vynálezom študentov Stafordskej Univerzity, Lawrencea Pagea a Sergeya Brina, ktorí v druhej polovici 90-tych rokov prišli s prelomovou myšlienkou vytvorenia nového webového vyhľadávača inšpirovanou nedostatkom relevantných informácií zobrazených medzi prvými výsledkami pri hľadaní. Založili spoločnosť Google, ktorej vyhľadávač sa od ostatných odlišoval predovšetkým faktorom nazývaný PageRank, ktorý umožnil usporadúvanie výsledkov podľa ich dôležitosti, ktorú vypočítal na základe odkazov smerujúcich na stránku, pričom bral do úvahy ich počet, hodnotenie a kvalitu odkazujúceho webu.

Takmer 15 rokov od začiatku prevádzky vyhľadávača Google sa faktor PageRank stále využíva, no nie je jediným faktorom, ktorý poradie zobrazených výsledkov ovplyvňuje.

Táto diplomová práca obsahuje zoznam ďalších približne dvesto signálov, ktoré majú na hodnotenie stránok vplyv. Približne 60 z týchto signálov sa nachádza v tabuľkách v prílohe – ide o náhodne vybrané webové stránky zobrazené medzi prvými desiatimi výsledkami pri vyhľadávaní vopred určeného kľúčového slova. Na základe tejto analýzy sme dospeli k záveru, že mnohé z týchto faktorov majú na hodnotenie vplyv – niektoré veľmi výrazný, niektoré minimálny. Aj táto váha vplyvu sa v priebehu času a vplyvom zmien vo webovom prostredí mení, Google zároveň skúma a vyhodnocuje stále nové faktory, ktoré by mohli pri hodnotení pomôcť.

V posledných dvoch rokoch, v čase aktualizácií Panda a Penguin, sa pozornosť v obrovskej miere sústredila na užívateľa a medzi hodnotiacimi faktormi stále lepšie postavenie získaval obsah stránky. Nie sú už podstatné len kľúčové slová, Google uprednostňuje obsah, ktorý je jedinečný, čerstvý a prínosný pre návštevníkov webov. Zároveň dáva priestor užívateľom – pri hodnotení stránok zohľadňuje ich aktivitu spojenú so sociálnymi sieťami, ako napríklad dávanie likeov stránkam, zdieľanie ich obsahu a pod.

Za kvalitné výsledky, ktoré vyhľadávač Google zobrazuje na prvých pozíciách, teda vďaka jedinečnému a neustále sa vyvíjajúcemu algoritmu, ktorého niektoré hodnotiace faktory síce poznáme, no podstata konečného hodnotenia pre nás zostáva neznáma.

## **ZOZNAM OBRÁZKOV V TEXTE**

|  |    |
|--|----|
| OBR. Č. 1: PROCES TVORBY ODPOVEDE NA UŽÍVATEĽSKÚ POŽIADAVKU .....                    | 18 |
| OBR.Č. 2: DOMOVSKÁ STRÁNKA VYHĽADÁVAČA GOOGLE PO JEHO SPUSTENÍ<br>(11.11.1998).....  | 20 |
| OBR. Č. 3: VÝVOJ HĽADANIA VÝRAZU „LAPTOPS“ OD ROKU 2004 PO SÚČASNOSŤ<br>.....        | 42 |
| OBR. Č. 4: STRÁNKA VÝSLEDKOV VYHĽADÁVAČA GOOGLE PRI HĽADANÍ HESLA<br>„LAPTOPS“ ..... | 43 |
| OBR. Č. 5: ČASOVÁ PRIAMKA AKTUALIZÁCIÍ GOOGLE .....                                  | 80 |

## ZOZNAM POUŽITÝCH ZDROJOV

Googlebot: Webmaster Tools Help. *Nástroje správcu webu* [online]. 2012, 2012-05-07 [cit. 2012-05-19]. Dostupné z: <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=182072>

Google Algorithm Change History. *SEO Software: SEOMoz* [online]. 2012 [cit. 2012-05-19]. Dostupné z: <http://www.seomoz.org/google-algorithm-change>

Top 15 Most Popular Search Engines. In: *EBizMBA: The eBusiness Knowledgebase* [online]. 2012 [cit. 2012-05-19]. Dostupné z: <http://www.ebizmba.com/articles/search-engines>

List of Google PageRank updates 2011. *Creative Drupal Web Design and SEO Company: French Website Designer France* [online]. 2011 [cit. 2012-05-19]. Dostupné z: <http://www.1websitedesigner.com/google-pagerank>

Google Toolbar PageRank Update List-History of PageRank Update(TBPR). *Internet Marketing Services* [online]. 2012 [cit. 2012-04-14]. Dostupné z: <http://www.searchenginegenie.com/google-toolbar-PageRank-update-list.html>

Google Launches The Google Toolbar. *Google: Inside Search* [online]. 2012 [cit. 2012-04-08]. Dostupné z: <http://www.google.com/press/pressrel/pressrelease42.html>

Google Pagerank Update History: May 2012 Update. ANDERSON HOB0, Shaun. *Hobo SEO Glasgow: Search Engine Optimisation Glasgow* [online]. 2012 [cit. 2012-04-14]. Dostupné z: <http://www.hobo-web.co.uk/google-pr-update/>

Gray PageRank Bar. In: *Slow Search Engine Optimization* [online]. 2007 [cit. 2012-04-14]. Dostupné z: <http://www.slowseo.com/articles/pagerank/grey-pagerank-bar.html>

Major Google Algorithm Updates of The Last Decade. *Internet Marketing Services* [online]. 2012 [cit. 2012-04-14]. Dostupné z: <http://www.usainternetmarketing.com/blog/google/major-google-algorithm-updates-of-the-last-decade-part-1/>

Google Updates Through The Years. In: *Orlando Website Design* [online]. 2011 [cit. 2012-04-14]. Dostupné z: <http://www.wsoaonline.com/google-updates-through-the-years/>

Google Updates Log. *Webmaster Social Network* [online]. 2011 [cit. 2012-04-20]. Dostupné z: <http://www.webmaisterpro.com/seo-guide/google-updates-log.html>

Google Dance Infographic finished. *Level343: San Francisco SEO - Organic International SEO & Copywriting* [online]. 2011 [cit. 2013-04-13]. Dostupné z: [http://level343.com/article\\_archive/wp-content/uploads/Google-Dance-Infographic-finished.png](http://level343.com/article_archive/wp-content/uploads/Google-Dance-Infographic-finished.png)

How Search Works. *Google: Inside Search* [online]. 2012 [cit. 2012-04-08]. Dostupné z: <http://www.google.com/insidesearch/howsearchworks/>

AVRACHENKOV, Konstantin a Nelly LITVAK. The Effect of New Links on Google Pagerank. *Stochastic Models: Estimation And Control: volume 1* [online]. New York: Academic Press, 2006, č. 22 [cit. 2012-07-03]. ISSN 1532-4214. Dostupné z: <http://doc.utwente.nl/63648/>

BESWICK, Matt. Google Algorithm Updates: All Time Top 10. *SEO Consultant - SEO Services: Matt Beswick* [online]. 2012 [cit. 2013-04-13]. Dostupné z: <http://www.mattbeswick.co.uk/google-algorithm-updates-all-time-top-10/>

BRIN, Sergey a Larry PAGE. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *The Stanford University InfoLab* [online]. 2008 [cit. 2013-01-17]. Dostupné z: <http://infolab.stanford.edu/~backrub/google.html>

CALISHAIN, Tara a Rael DORNFEST. *Google hacks: 100 industrial-strength tips*. Sebastopol, CA: O'Reilly, 2003, xxi, 329 p. ISBN 05-960-0447-8.

CONTI, Greg. *Googling security: how much does Google know about you?*. Upper Saddle River, NJ: Addison-Wesley, c2009, xxi, 332 p. ISBN 978-032-1518-668.

CRAVEN, Phil. Pagerank Explained: Google's PageRank and how to make the most of it. *Search engine optimization services and articles, inc.: PageRank Explained, search engine optimization forum. UK based, worldwide clients*. [online]. [2003] [cit. 2012-07-02]. Dostupné z: <http://www.webworkshop.net/pagerank.html>

CUTTS, Matt. Explaining algorithm updates and data refreshes. In: CUTTS, Matt. *Matt Cutts: Gadgets, Google, and SEO* [online]. 2006 [cit. 2012-04-15]. Dostupné z: <http://www.mattcutts.com/blog/explaining-algorithm-updates-and-data-refreshes/>

DEAN, Brian. Google's 200 Ranking Factors: The Complete List. In: *Backlinko: Link Building for SEO Domination* [online]. 2013 [cit. 2013-04-23]. Dostupné z: <http://backlinko.com/google-ranking-factors>

DHOLAKIYA, Pratik. Google Search Algorithm Updates in 2012. In: *E2M Solutions: Full Service Professional Internet Marketing Company* [online]. 2013 [cit. 2013-04-16]. Dostupné z: <http://www.e2msolutions.com/blog/google-search-algorithm-updates-in-2012-infographic/>

GERON, Tomio. Google's 4 Biggest Technical Challenges, According To Search Guru Amit Singhal. In: *Forbes.com: Information for the World's Business Leaders* [online]. 2013 [cit. 2013-04-29]. Dostupné z: <http://www.forbes.com/sites/tomiogeron/2013/03/10/googles-4-biggest-technical-challenges-according-to-search-guru-amit-singhal-sxsw/>

GRAPPONE, Jennifer a Gradiva COUZIN. *Search engine optimization: an hour a day*. Indianapolis: Wiley, c2006, xviii, 318 s. ISBN 04-717-8753-1.

GUNTER, Barrie, Ian ROWLANDS a David NICHOLAS. *The Google generation: are ICT innovations changing information-seeking behaviour?*. Oxford: Chandos Pub., 2009, xi, 207 p. ISBN 978-184-3345-572.

GUPTA, Deepak. A Timeline of Google Algorithm Changes & Updates - Panda 3.4 'Rolling Out Now'. *DiTii.com - All About Technology Blog RSS Feed* [online]. 2012 [cit. 2013-03-09]. Dostupné z: <http://www.ditii.com/2012/03/26/timeline-google-algorithm-updates-panda-3-4-rolling/>

HOCK, Randolph. *The extreme searcher's Internet handbook: a guide for the serious searcher*. 2nd ed. Medford, N.J.: CyberAge Books, c2007, xxvii, 326 p. ISBN 09-109-6576-5.

HOUSER, Pavel. Scienceworld: Matematika za Google PageRankem není tajná. *Science World.cz: Novinky ze světa vědy a techniky: technologie, neživá příroda, člověk, biologie* [online]. 2008 [cit. 2012-07-07]. Dostupné z: <http://scienceworld.cz/technologie/matematika-za-google-pagerankem-neni-tajna>

JANOUSH, Marek. Google chce rozumět tomu, co hledáte. In: *Lupa.cz: server o českém Internetu* [online]. 2012 [cit. 2013-04-26]. Dostupné z: <http://www.lupa.cz/clanky/google-knowledge-graph/>

KLUVER, Casey. A Visual Look at 2011 Google Algorithm Changes. *SEO Software. Simplified. / SEOMoz* [online]. 2012 [cit. 2013-03-17]. Dostupné z: <http://www.seomoz.org/ugc/a-visual-look-at-2011-google-algorithm-changes>

LONG, Johnny. *Google hacking*. Vyd. 1. Zoner Press: Brno, 2005, 472 s. ISBN 80-868-1531-5.



MAKULOVÁ, Soňa. *Vyhľadávanie informácií v internete: Problémy, východiská, postupy*. 1. vyd. Rudolf Bista. Jozef Mlacek. Bratislava: EL&T, 2002, 376 s. ISBN 80-88812-16-X.

ŠUBR, Jiří. Blog Jiří Šubr: Internetový marketing a byznys. In: ŠUBR. *Blog Jiří Šubr: Internetový marketing a byznys* [online]. 2013 [cit. 2013-04-28]. Dostupné z: <http://blog.subrjiri.cz/otazky-a-odpovedi-webtrh-1/>

WISE, David A a Mark MALSEED. *The Google story*. Delta trade pbk. ed. New York: Delta Trade Paperbacks, 2006c2005, x, 326 p. ISBN 05-533-8366-3.

WILLS, Rebecca S. Google's PageRank: The Math Behind the Search Engine. In: *CEMS Home: University of Vermont* [online]. 2006 [cit. 2012-07-03]. Dostupné z: [http://www.cems.uvm.edu/~tlakoba/AppliedUGMath/other\\_Google/Wills.pdf](http://www.cems.uvm.edu/~tlakoba/AppliedUGMath/other_Google/Wills.pdf)

# ZOZNAM PRÍLOH

## Slovník základných pojmov

## Zoznam použitých analytických nástrojov

### Tabuľky:

|                       |  |
|-----------------------|--|
| <b>Tabuľka č. 1:</b>  | Analyzované stránky                                  |
| <b>Tabuľka č. 2:</b>  | Faktory na úrovni domény                             |
| <b>Tabuľka č. 3:</b>  | Faktory na úrovni stránky                            |
| <b>Tabuľka č. 4:</b>  | Faktory na úrovni stránky (pokračovanie)             |
| <b>Tabuľka č. 5:</b>  | Faktory na úrovni webového sídla                     |
| <b>Tabuľka č. 6:</b>  | Faktory súvisiace so spätnými odkazmi                |
| <b>Tabuľka č. 7:</b>  | Faktory súvisiace so spätnými odkazmi (pokračovanie) |
| <b>Tabuľka č. 8:</b>  | Faktory súvisiace so spätnými odkazmi (pokračovanie) |
| <b>Tabuľka č. 9:</b>  | Špeciálne pravidlá algoritmu                         |
| <b>Tabuľka č. 10:</b> | Sociálne signály                                     |
| <b>Tabuľka č. 11:</b> | Signály týkajúce sa značky                           |

### Obrázky:

|                    |   |
|--------------------|---|
| <b>Obr. č. 6:</b>  | Rýchlosť rastu počtu odkazov na web bestbuy.com |
| <b>Obr. č. 7:</b>  | Rýchlosť rastu počtu odkazov na web dell.com    |
| <b>Obr. č. 8:</b>  | Rýchlosť rastu počtu odkazov na web lenovo.com  |
| <b>Obr. č. 9:</b>  | Rýchlosť rastu počtu odkazov na web pcworld.com |
| <b>Obr. č. 10:</b> | Rýchlosť rastu počtu odkazov na web hp.com      |
| <b>Obr. č. 11:</b> | Zobrazenie bestbuy.com v Google správach        |
| <b>Obr. č. 12:</b> | Zobrazenie dell.com v Google správach           |
| <b>Obr. č. 13:</b> | Zobrazenie lenovo.com v Google správach         |
| <b>Obr. č. 14:</b> | Zobrazenie pcworld.com v Google správach        |
| <b>Obr. č. 15:</b> | Zobrazenie hp.com v Google správach             |

# PRÍLOHY

## Slovník základných pojmov

ANCHOR TEXT – viditeľný text, na ktorý sa dá kliknúť hypertexte. Slová v anchor texte môžu určiť ranking stránky určený vyhľadávacími zariadeniami

AUTORITATÍVNE WEB SÍDLA (Authority Sites) – stránky ktorým Google pripisuje väčšiu váhu, ako Wikipedia, .gov, .edu stránky

Autorstvo Google+ (Google+ Authorship) – spájanie obsahu uverejneného na určitej web stránke s Google+ profilom

BAD NEIGHBORHOODS - spam, stránky používajúce zakázané techniky a pod.

BARELY – (barrels), súbory hitov pre forward index

CRAWLING – prezeranie - proces, ktorým Googlebot objavuje nové a aktualizované stránky, aby ich mohol pridať do Google indexu.

CTR - (Click Through Rate), miera prekliku, pomer medzi počtom zobrazení reklamnej plochy a počtom kliknutí na reklamu

DESCRIPTION TAG – časť kódy ktorá obsahuje popis a stručné informácie o stránke

DIGITAL MILLENNIUM COPYRIGHT ACT (DMCA) - zákon dopĺňujúci americké autorské právo Copyright Act

DIRECTORY.MOZILLA.ORG (DMOZ) - verejne dostupný register odkazov v rámci WWW (World Wide Web)

DISAVOW TOOL - nástroj na oznámenie odkazov, ktoré by nemali byť uznané

DMCA - Digital Millennium Copyright Act, obdoba Autorského práva v Spojených štátoch

DOCID – unikátny identifikátor dokumentu

DOMÉNA .COM - generická doména najvyššieho radu, povodne určená pre komerčné účely, v súčasnosti však jej použitie nie je obmedzené

DUMPLEXICON – program, ktorý vytvára lexikón pre použité vyhľadávačom

EIGENVECTOR - vlastný vektor, ktorý je nenulový a jeho smer sa pri transformácii nemení

EXACT MATCH DOMAIN (EMD) - presná zhoda názvu domény s vyhľadávacím dotazom

FACEBOOK - sociálna sieť založená vo februári 2004 študentmi z Harvardskej Univerzity

FOLLOWER - označenie pre fanúšika stránky na Twitteri

FORWARD INDEX – obsahuje zoznam slov pre každý dokument

GOOGLE DANCE – mesačná aktualizácia indexu, zmeny kombinácií algoritmov

GOOGLE NEWS BOX - okno s novinkami Google

GOOGLE SANDBOX – zoznam stránok, väčšinou nových, ktoré sa dočasne nezobrazujú na pozíciách zodpovedajúcich ich PageRanku

GOOGLE TOOLBAR - vyhľadávanie Google z lišty webového prehľadávača

H1, H2, H3 TAGY – nadpisy prvej, druhej, tretej úrovne

HITY – (hits), hity zaznamenávajú slovo, jeho pozíciu v dokumente, približnú veľkosť fontu a pozíciu veľkých písmen

HODNOTENIE +1 - obdobná funkcia Like na Facebooku používaná v rámci sociálnej siete Google Plus

HOME PAGE – domovská stránka webu

HTML ERROR – chyba v HTML kóde

HTML ROZHRANIE - (Hypertext Markup Language based interface), je značkový jazyk určený na vytváranie webových stránok a iných informácií zobraziteľných vo webovom prehliadači.

HUB PAGES - top stránky na určitú tému

IMAGE RESULTS – výsledky vyhľadávania v obrázkoch Google

INDEX – usporiadaný zoznam objektov, názvov, tém apod. spolu s informáciou, kde možno objekty nájsť; databáza web stránok prispôbena na vyhľadávanie

INDEXAČNÝ PROGRAM – program, ktorý spracúva informácie získané robotmi

INDEXER (INDEXÁTOR) – spracováva súbory a ukladá ich do databázy, číta obsah repozitára, dekomprimuje a analyzuje dokumenty

INDEXOVANIE - spracovávanie súborov a ich ukladanie do databázy

IP ADRESA – unikátna adresa zariadenia v sieti

KANONICKÁ STRÁNKA – stránka s rovnakým obsahom ale iným usporiadaním

KEYWORD (KW) – kľúčové slovo

KRUHY Google+ - (Circles) – filtre používané v Google + umožňujúce zdieľanie obsahu len s vybranými používateľmi

LIKE - funkcia sociálnej siete Facebook, ktorá umožňuje používateľom oceniť publikovaný obsah.

LINKEDIN - sociálna sieť určená primárne ľuďom v odborných profesiách

METATAGY - informácie o stránke, ktoré oznamujú prehliadačom údaje o stránke; kto je autorom stránky, v akom kóde je napísaná, aký je názov stránky, apod.

MIERA OPUSTENIA - (Bounce Rate), pomer medzi počtom návštevníkov, ktorí ihneď odišli, tj. nikam neklikli, a celkovým počtom návštevníkov v percentách

MIKROFORMÁT – množina jednoduchých, otvorených dátových formátov, vybudovaných na existujúcich a rozšírených štandardoch

MININAVIGÁCIA - (Omrvinková navigácia, Breadcrumb Navigation), navigačná pomôcka používaná v rámci užívateľského rozhrania, umožňuje užívateľovi mať prehľad o svojej aktuálnej pozícii v rámci programu alebo dokumentu

NEWS FEED - stránky s novinkami

NOFOLLOW ODKAZY - prostredníctvom atribútu Nofollow dávajú správcovia webov príkaz, aby Google odkazy na danej stránke ďalej nesledoval

ODKAZOVÉ FARMY – katalóg, ktorý má vysoký PageRank, pri zaregistrovaní stránky na tento katalóg získa stránka vyšší PageRank

ODKAZY ZA PROVÍZIU - (Affiliate Links), odkazy na partnerské stránky

PARKOVANÉ DOMÉNY - (Parked domains), zaregistrovanie názvu domény bez toho aby bola využívaná

PINTEREST - sociálna sieť založená na princípe nástenky pre zdieľanie fotografií, ktoré sa nazývajú označením PIN

PRESMEROVANIE 301 – statické presmerovanie na inú časť stránky

RANKING – rating, hodnotenie podľa vopred daných vzorcov, vyjadruje vzťah medzi skupinou objektov tak, že jeden je hodnotený nižšie, rovnako, príp. vyššie ako druhý

REL=CANONICAL – časť kódu v odkaze, ktorá umožňuje smerovanie na kanonickú stránku

RICH SITE SUMMARY (RSS) - formát súboru (Rich Site Summary) je skupina formátov určených na čítanie noviniek na webových stránkach a všeobecne o syndikácii obsahu

ROBOT (SPIDER, CRAWLER) – softvérová aplikácia, ktorá vykonáva automatizované úlohy v prostredí internetu, zväčša prehľadávanie za účelom indexovania

RODINY WEBOVÝCH SÍDEL - (web site families, link wheels), súbor linkov smerujúcich od jednej web stránky na ďalšiu až k cieľovej stránke, za účelom zvýšenia jej PageRanku

SEARCH QUERY - hľadané heslo, vyhľadávacia požiadavka, kľúčové slovo zadané do okna vyhľadávača

SEO – (Search Engine Optimization), optimalizácia obsahu stránok pre vyhľadávače

SHOPPING RESULTS - Výsledky z oblasti nakupovania

SITEMAP – mapa webu

SORTER (TRIEDIČ) – triedi dokumenty pre indexáciu

SPLOG - označenie spamových blogov

SSL CERTIFIKÁT - digitálne podpísaný elektronický dokument, verejný šifrovací kľúč

STRÁNKA 404 – neexistujúca stránka

SYNDIKOVANÝ OBSAH – informácie poskytované prostredníctvom služieb na internete pre použitie na webových stránkach

TITLE TAG - Nadpis stránky, jeden z najdôležitejších údajov o stránke, v zdrojovom kóde je označený tagom <title>

TLD – (Top Level Domain), doména najvyššieho rádu, najvyššia úroveň domény

TWEET - označenie krátkej správy (max. 140 znakov) odoslanej prostredníctvom sociálnej siete Twitter

TWITTER - online služba, sociálna sieť pre odosielanie textových správ

ÚLOŽNÝ SERVER - (store server) počítač pripojený k sieti, ktorého primárnou úlohou je poskytovať prístup k zdieľaným dátam zariadeniam pripojeným v sieti

UNIFORM RESOURCE LOCATOR (URL) - adresa lokalizácie na internete

UPDATE - aktualizácia

URL ADRESA - jednotný vyhľadávač zdrojov (Uniform Resource Locator), skratka URL, je univerzálny formát mien používaný na označenie zdroja na internete

URL DEKÓDER – online nástroj ktorý konvertuje určité znaky do URL a nahrádza ich jednou alebo viacerými trojicami znakov obsahujúcimi znamienko "%" a dve šestnástkové čísla, ktoré predstavujú numerickú hodnotu znaku.

URL SERVER – úložisko URL adries

WC3 VALIDÁCIA – overenie či stránka spĺňa štandard WC3

WEBMASTER – správca webu

WHOIS - databáza údajov o majiteľoch internetových domén a IP adries

WORLD WIDE WEB (WWW) – celosvetová sieť, systém prepojených hypertextových dokumentov dostupných prostredníctvom internetu

ZLOMENÉ ODKAZY – (broken links), nefungujúce odkazy odkazujúce na neexistujúcu stránku

## **Zoznam použitých analytických nástrojov**

### **Online**

<http://www.seologs.com/dns/domain-check.html>

<http://www.whois.com>

<http://www.seocentro.com/tools/seo/keyword-density.html>

<https://developers.google.com/speed/pagespeed/insights>

<https://ahrefs.com/index.php>

<https://www.majesticseo.com>

<http://www.seomoz.org/tools>

<http://pagerankchecker.com>

<http://www.seocentro.com>

<http://www.seomastering.com/trust-rank-checker.php>

<http://www.responsivedesignchecker.com/>

<https://adwords.google.com>

<http://www.google.com/webmasters/tools/>

### **Desktop**

Link Assistant

Website Auditor

## Tabuľky a obrázky

| OZNAČENIE<br>ANALYZOVANEJ<br>STRÁNKY | URL   | POZÍCIA<br>NA KLÚČOVÉ<br>SLOVO "LAPTOPS" | DOMÉNA      | PAGERANK<br>DOMÉNY |
|--------------------------------------|---|--|-------------|--------------------|
| Stránka A                            | http://www.bestbuy.com/site/Computers-PCs/Laptop-Computers/abcat0502000.c?id=abcat0502000 | 1.                                       | bestbuy.com | 7                  |
| Stránka B                            | http://www.dell.com/us/p/laptops  | 3.                                       | dell.com    | 8                  |
| Stránka C                            | http://www.lenovo.com/products/us/laptop/   | 6.                                       | lenovo.com  | 8                  |
| Stránka D                            | http://www.pcworld.com/category/laptop-computers  | 8.                                       | pcworld.com | N/A                |
| Stránka E                            | http://www.hp.com/country/us/en/pr odserv/laptops.html                                    | 10.                                      | hp.com      | 9                  |

Tabuľka č. 1: Analyzované stránky

| STRÁNKA | VYTVORENIE<br>DOMÉNY | VEK DOMÉNY <sup>3</sup>     | PLATNOSŤ<br>DOMÉNY VYPRŠÍ <sup>4</sup> | ZÁZNAM WHOIS <sup>5</sup> | POSLEDNÝ<br>UPDATE<br>ZÁZNAMU | KW "LAPTOP<br>S" V NÁZVE<br>DOMÉNY |
|---------|----------------------|-----------------------------|--|---------------------------|-------------------------------|------------------------------------|
| A       | 2.3.1994             | 19 rokov 1 mesiac<br>4      | 3.3.2020<br>2                          | áno, verejný              | 31.12.2012                    | nie                                |
| B       | 22.11.1988           | 24 rokov<br>5 mesiacov<br>2 | 21.11.2015<br>4                        | áno, verejný              | 21.11.2011                    | nie                                |
| C       | 5.9.2002             | 10 rokov<br>7 mesiacov<br>5 | 6.9.2016<br>3                          | áno, verejný              | 20.6.2011                     | nie                                |
| D       | 24.4.1992            | 21 rokov<br>3               | 24.4.2020<br>1                         | áno, verejný              | 16.4.2013                     | nie                                |
| E       | 3.3.1986             | 27 rokov<br>2 mesiace<br>1  | 3.3.2014<br>5                          | áno, verejný              | 31.1.2013                     | nie                                |

Tabuľka č. 2: Faktory na úrovni domény

<sup>3</sup> staršie domény - dôveryhodnejšie (nie príliš dôležitý faktor)

<sup>4</sup> domény zaplatené na niekoľko rokov dopredu, môžu byť vnímané ako legitímne, čo pozitívne ovplyvní ich hodnotenie

<sup>5</sup> zapnuté skrývanie súkromných informácií vo Whois môže evokovať, že sa Webmaster snaží niečo zatajiť



| STRÁNKA | KLÚČOVÉ SLOVO V <TITLE> <sup>6</sup>  | KLÚČOVÉ SLOVO V <DESCRIPTION>  | KLÚČOVÉ SLOVO V H1             | DĚLKA OBSAHU <sup>7</sup> | CELKOVÝ POČET KLÚČOVÝCH SLOV | HUSTOTA SLOVA "LAPTOPS" <sup>8</sup> | RÝCHLOŠŤ NAČÍTANIA STRÁNKY (GOOGLE SCORE) |
|---------|---|--|--------------------------------|---------------------------|------------------------------|--------------------------------------|---|
| A       | Laptops & Notebook Computers - Best Buy   | Read customer reviews, compare prices and buy online at Best Buy.  | Laptops                        | 2197 slov                 | 1835                         | 24 výskytov (1,31%)                  | 70/100                                    |
|         | 1   | 5  |                                | 1                         |                              |                                      | 5   |
| B       | Laptops, Ultrabooks™ & Tablets - New Dell Laptop Computers for Sale   Dell                        | Laptop & Notebook Computer Sales from the Official Dell Site. Dell offers everyday, performance, ultra-thin and gaming ready laptops. Build and ship yours today!  | Laptops, Ultrabooks™ & Tablets | 1248 slov                 | 889                          | 10 výskytov (1,12%)                  | 81/100                                    |
|         | 1   | 4  |                                | 2                         |                              |                                      | 2   |
| C       | Laptop and Ultrabook PCs - Best, Most Reliable and Powerful Computers in Industry   Lenovo   (US) | See how Lenovo Ultrabooks and laptops are designed for ultraportability and extra long battery life -- some systems run up to 30 hours. We keep you unplugged whether you're at work or play.  | Lenovo Laptops & Ultrabooks    | 781 slov                  | 2018                         | 27 výskytov (1,34%)                  | 79/100                                    |
|         | 5   | 3  |                                | 4                         |                              |                                      | 3   |
| D       | Laptops reviews, how to advice, and news  | Get the latest Laptops reviews, Laptops buying guides, and Laptops prices from the knowledgeable experts at PCWorld  | Laptops                        | 788 slov                  | 947                          | 11 výskytov (1,16%)                  | 88/100                                    |
|         | 1   | 1  |                                | 3                         |                              |                                      | 1   |
| E       | Laptops and Notebook Computers   HP Laptop Computers  | Find the HP laptop just for you. As the #1 provider of laptops in the world, HP offers affordable and powerful laptops and notebook computers. Compare and shop. Buy with confidence from the #1 provider of PCs to worldwide businesses | Laptops and Notebook Computers | 536 slov                  | 1071                         | 20 výskytov (1,87%)                  | 77/100                                    |
|         | 1   | 2  |                                | 5                         |                              |                                      | 4   |

Tabuľka č. 3: Faktory na úrovni stránky

<sup>6</sup> „Title“ tag začínajúci kľúčovým slovom má pravdepodobne väčšiu váhu, ako keď sa kľúčové slovo nachádza v strede alebo na konci

<sup>7</sup> stránky s väčším počtom slov v obsahu majú potenciál objavovať sa vo výsledkoch vyššie

<sup>8</sup> v poslednej dobe Google hustotu kľúčových slov využíva predovšetkým na určovanie témy webu, ak je príliš vysoká, na hodnotenie stránky má negatívny dopad

| STRÁNKA | Optimalizácia obrázkov <sup>9</sup> | Autorita stránky | Zlomené odkazy <sup>10</sup> | Kvalita kódu (HTML error/W3C validácia)             | PR stránky | príliš dlhé URL (viac než 115 znakov) <sup>11</sup> | Kľúčové slovo v URL | Odrážky a číslované zoznamy <sup>12</sup> | Priorita mapy webu <sup>13</sup> |
|---------|-------------------------------------|------------------|------------------------------|---|------------|---|---------------------|---|----------------------------------|
| A       | používa alt tagy                    | 81               | 255 (5,5%)                   | 10 stránok, 4 723 W3C errors, 3 911 varovaní        | 6          | 0 stránok   | nie                 | používa                                   | áno                              |
|         | 2                                   | 2                |                              |   |            |   |                     |   |                                  |
| B       | používa alt aj title tagy           | 88               | 224 (11%)                    | 220 stránok, 76 631 W3C errors, 47 334 varovaní     | 6          | 13 stránok  | áno                 | používa                                   | áno                              |
|         | 1                                   | 1                |                              |   |            |   |                     |   |                                  |
| C       | používa alt tagy                    | 67               | 4298 (11,7%)                 | 628 stránok, 32 548 W3C errors, 8 762 varovaní      | 5          | 0 stránok   | nie                 | používa                                   | áno                              |
|         | 2                                   | 3                |                              |   |            |   |                     |   |                                  |
| D       | nepoužíva                           | 59               | 401 (0,2%)                   | 1 539 stránok, 200 737 W3C errors, 457 273 varovaní | N/A        | 67 stránok  | nie                 | používa                                   | áno                              |
|         | 3                                   | 5                |                              |   |            |   |                     |   |                                  |
| E       | používa alt tagy                    | 61               | 10 784 (13,5%)               | 708 stránok, 70053 W3C errors, 31 232 varovaní      | 0          | 134 stránok   | áno                 | používa                                   | áno                              |
|         | 2                                   | 4                |                              |   |            |   |                     |   |                                  |

Tabuľka č. 4: Faktory na úrovni stránky (pokračovanie)

<sup>9</sup> poskytujú podstatné signály vyhľadávateľom cez názov súboru alt text, nadpis (title), popis (description) a titulok (caption)

<sup>10</sup> môžu ukazovať na zanedbanú alebo nepoužívanú stránku

<sup>11</sup> môžu znížiť viditeľnosť stránok

<sup>12</sup> zvyšujú prehľadnosť textu pre návštevníkov, a preto ich kladne hodnotí aj Google

<sup>13</sup> použitie sitemap.xml ukáže váhu jednotlivých stránok webu

| STRÁNKA | STRÁNKA „KONTAKTUJTE NÁS“ <sup>14</sup> | TRUST RANK <sup>15</sup> | POČET STRÁNOK WEBU (PODĽA INDEXU GOOGLE) | LOKALIZÁCIA SERVERA <sup>16</sup> | VŠEOBECNÉ PODMIENKY A OCHRANA SÚKROMIA <sup>17</sup> | DUPLICITNÝ OBSAH WEBU <sup>18</sup> | YOUTUBE <sup>19</sup>                            | POUŽÍVANIE GOOGLE ANALYTICS A GOOGLE WEBMASTER TOOLS <sup>20</sup> |
|---------|---|--------------------------|--|-----------------------------------|--|-------------------------------------|--|--|
| A       | áno                                     | 9,82/10                  | 28 900 000                               | us                                | áno  | nie                                 | youtube.com/user/bestbuy<br>5 776 odoberateľov   | nie  |
|         |   | 2                        |  |                                   |  |                                     | 5  |  |
| B       | áno                                     | 10.10                    | 40 900 000                               | us                                | áno  | nie                                 | youtube.com/user/DellVlog<br>15 701 odoberateľov | nie  |
|         |   | 1                        |  |                                   |  |                                     | 2  |  |
| C       | áno                                     | 9,37/10                  | 13 200                                   | us                                | áno  | nie                                 | youtube.com/lenovo<br>17 180 odoberateľov        | nie  |
|         |   | 3                        |  |                                   |  |                                     | 1  |  |
| D       | áno                                     | 4,65/10                  | 1 840 000                                | us                                | áno  | nie                                 | youtube.com/PCWorld<br>13 341 odoberateľov       | áno  |
|         |   | 4                        |  |                                   |  |                                     | 3  |  |
| E       | áno                                     | 10.10                    | 398 000                                  | us                                | áno  | nie                                 | youtube.com/user/HP<br>13 053 odoberateľov       | nie  |
|         |   | 1                        |  |                                   |  |                                     | 4  |  |

Tabuľka č. 5: Faktory na úrovni webového sídla

<sup>14</sup> stránky s kontaktnými informáciami môžu byť uprednostnené, o to viac ak sa zhodujú s Whols

<sup>15</sup> TrustRank je vyjadrenie pre dôveryhodnosť domény - veľmi vplyvný faktor, Google hodnotí prostredníctvom Trust Ranku, ktorý ovplyvňuje mnoho faktorov, a na jeho základe odlišuje užitočné stránky od spamu

<sup>16</sup> významná pri geograficky obmedzených hľadániach

<sup>17</sup> Terms of Service) a ochrana súkromia (Privacy Pages)

<sup>18</sup> duplicitné stránky či meta informácie jedného webu prispievajú k zápornému hodnoteniu

<sup>19</sup> je dokázané, že videa z YouTube môžu byť vo výsledkoch na SERP preferované – zobrazenie na vrchných pozíciách s náhľadom

<sup>20</sup> odhaduje sa pozitívny dopad na pozície webu vďaka presným dátam, ktoré zaznamenávajú a sprístupňujú Google

| STRÁNKA | POČET DOMÉN ODKAZUJÚCICH NA WEB <sup>21</sup> | POČET DOMÉN ODKAZUJÚCICH NA STRÁNKU | POČET ODKAZUJÚCICH STRÁNOK NA WEB <sup>22</sup> | POČET ODKAZUJÚCICH STRÁNOK NA STRÁNKU | POČET INDEXOVANÝCH STRÁNOK WEBU | POČET ODKAZOV ZO SEPARÁTNÝCH IP ADRES TRIEDY C NA WEB | IP ADRESY SMERUJÚCE NA WEB | POČET ODKAZOV ZO SEPARÁTNÝCH IP ADRES TRIEDY C NA STRÁNKU |
|---------|---|-------------------------------------|---|---------------------------------------|---------------------------------|---|----------------------------|---|
| A       | 81 890  | 373                                 | 2 823 795                                       | 2 259                                 | 761 835                         | 27 775  | 52 154                     | 298   |
|         | 4   | 2                                   | 5   | 3                                     | 3                               | 4   | 4                          | 2   |
| B       | 161 317                                       | 828                                 | 10 589 612                                      | 4 392                                 | 1 162 062                       | 53 565  | 103 098                    | 490   |
|         | 2   | 1                                   | 2   | 2                                     | 2                               | 2   | 2                          | 1   |
| C       | 55 085  | 52                                  | 3 417 167                                       | 294                                   | 436 457                         | 25 050  | 39 609                     | 48  |
|         | 5   | 4                                   | 4   | 4                                     | 4                               | 5   | 5                          | 4   |
| D       | 148 411                                       | 116                                 | 6 688 205                                       | 16 004                                | 369 980                         | 44 140  | 88 756                     | 106   |
|         | 3   | 3                                   | 3   | 1                                     | 5                               | 3   | 3                          | 3   |
| E       | 271 727                                       | 48                                  | 19 877 227                                      | 126                                   | 2 687 322                       | 78 439  | 162 488                    | 44  |
|         | 1   | 5                                   | 1   | 5                                     | 1                               | 1   | 1                          | 5   |

Tabuľka č. 6: Faktory súvisiace so spätnými odkazmi

| STRÁNKA | IP ADRESY SMERUJÚCE NA STRÁNKU | POČET EXTERNÝCH ODKAZOV DOMOVSKÉJ STRÁNKY | POČET INTERNÝCH ODKAZOV DOMOVSKÉJ STRÁNKY | ODKAZY NA WEB Z DOMÉN .EDU A .GOV <sup>23</sup> | NOFOLLOW ODKAZY | NAJČASTEJŠIE POUŽÍVANÉ ANCHOR TEXTY SPÄTNÝCH ODKAZOV  |
|---------|--------------------------------|---|---|---|-----------------|---|
| A       | 334                            | 19  | 266                                       | Gov: 90<br>Edu: 389                             | 255 (6,3%)      | 17% odkazov: "best buy"<br>9% odkazov: žiadny text<br>7% odkazov: "bestbuy.com"                                 |
|         | 2                              | 3   | 1   | 4   |                 |   |
| B       | 559                            | 23  | 90  | Gov: 219<br>Edu: 1722                           | 0               | 9% odkazov: žiadny text<br>6% odkazov: "dell"<br>2% odkazov: "about our ads"                                    |
|         | 1                              | 2   | 3   | 2   |                 |   |
| C       | 51                             | 6   | 12  | Gov: 78<br>Edu: 390                             | 2 (0,1%)        | 10% odkazov: žiadny text<br>8% odkazov: "lenovo"<br>4% odkazov: "联想集团"<br>(v preklade z čínštiny: Lenovo Group) |
|         | 4                              | 4   | 5   | 5   |                 |   |
| D       | 111                            | 27  | 16  | Gov: 133<br>Edu: 1075                           | 7 255 (3,1%)    | 19% odkazov: "pc world"<br>4% odkazov: "pcworld"<br>2% odkazov: "pcworld usa"                                   |
|         | 3                              | 1   | 4   | 3   |                 |   |
| E       | 47                             | 3   | 138                                       | Gov: 407<br>Edu: 2417                           | 21 (0,2%)       | 29% odkazov: "hewlett-packard"<br>15% odkazov: žiadny text<br>8% odkazov: "hp"                                  |
|         | 5                              | 5   | 2   | 1   |                 |   |

<sup>21</sup> Jeden z najdôležitejších hodnotiacich faktorov algoritmu Google

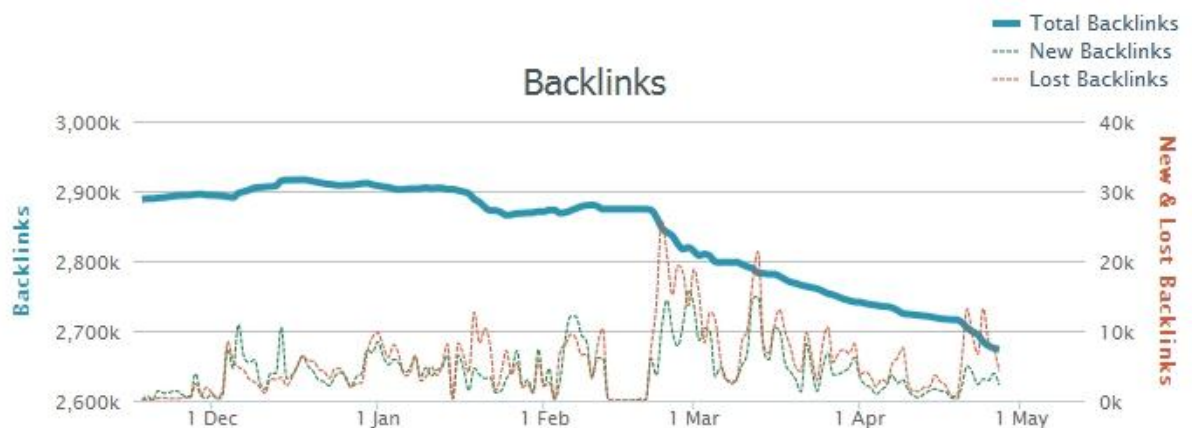
<sup>22</sup> bez ohľadu na to, či sú na spoločnej doméne

<sup>23</sup> Podľa Matta Cuttsa tento faktor nemá na hodnotenie vplyv

Tabuľka č. 7: Faktory súvisiace so spätnými odkazmi (pokračovanie)

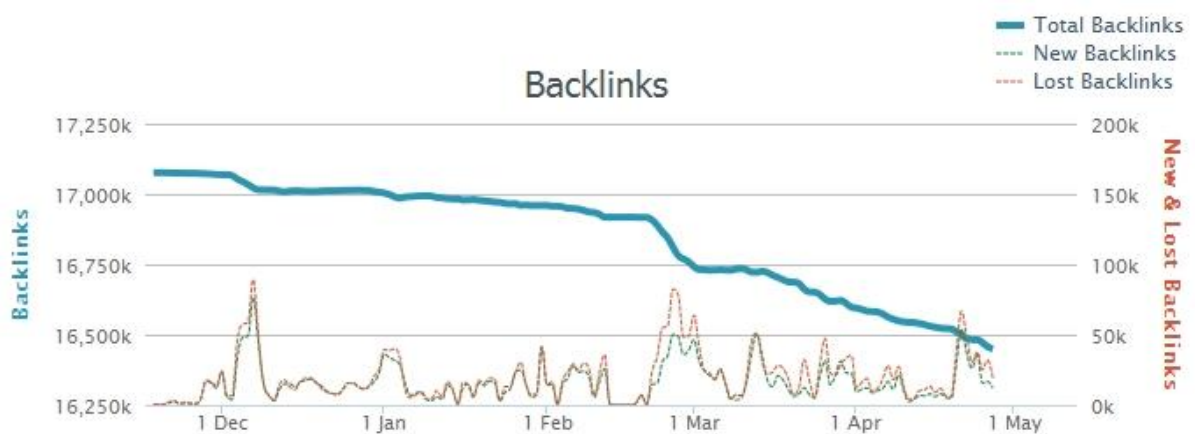
| STRÁNKA | RÝCHLOSŤ RASTU<br>POČTU ODKAZOV<br>STRÁNKY | ODKAZ<br>NA WEB<br>Z WIKIPEDIE | ODKAZY<br>Z PRESMEROVANIA<br>301 | UMIESTNENIE<br>V ZOZNAM DMOZ | UMIESTNENIE<br>V REGISTR<br>YAHOO! |
|---------|--|--------------------------------|----------------------------------|------------------------------|------------------------------------|
| A       | obr. č. 6                                  | existuje                       | 0 odkazov                        | áno                          | áno                                |
|         |  |                                | 1                                |                              |                                    |
| B       | obr. č. 7                                  | existuje                       | 147 (16,5%) odkazov              | áno                          | áno                                |
|         |  |                                | 4                                |                              |                                    |
| C       | obr. č. 8                                  | existuje                       | 314 (48,6%) odkazov              | áno                          | nie                                |
|         |  |                                | 3                                |                              |                                    |
| D       | obr. č. 9                                  | existuje                       | 128 (8,3%) odkazov               | áno                          | áno                                |
|         |  |                                | 2                                |                              |                                    |
| E       | obr. č. 10                                 | existuje                       | 461 (63%) odkazov                | áno                          | áno                                |
|         |  |                                | 5                                |                              |                                    |

Tabuľka č. 8: Faktory súvisiace so spätnými odkazmi (pokračovanie)



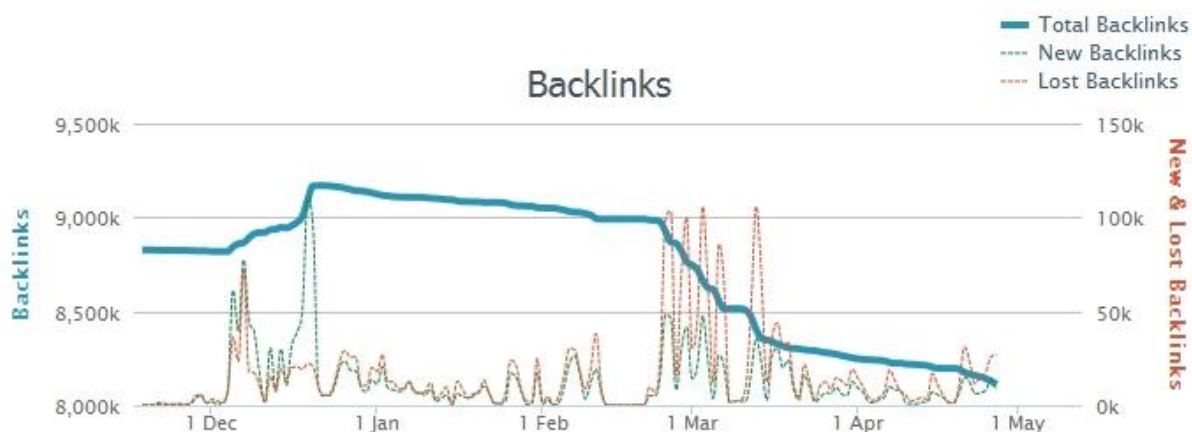
Obr. č. 6: Rýchlosť rastu počtu odkazov na web bestbuy.com

(Zdroj: <https://ahrefs.com>)



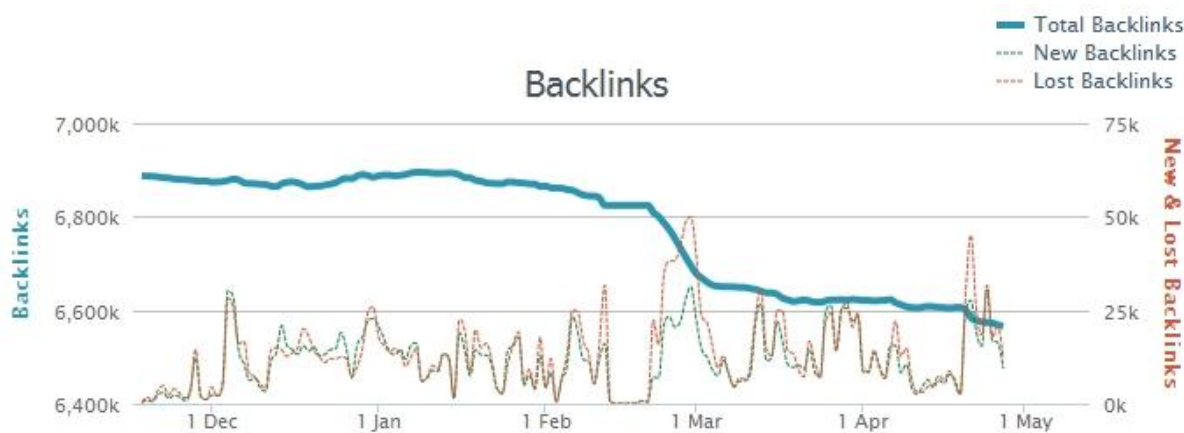
Obr. č. 7: Rýchlosť rastu počtu odkazov na web dell.com

(Zdroj: <https://ahrefs.com>)



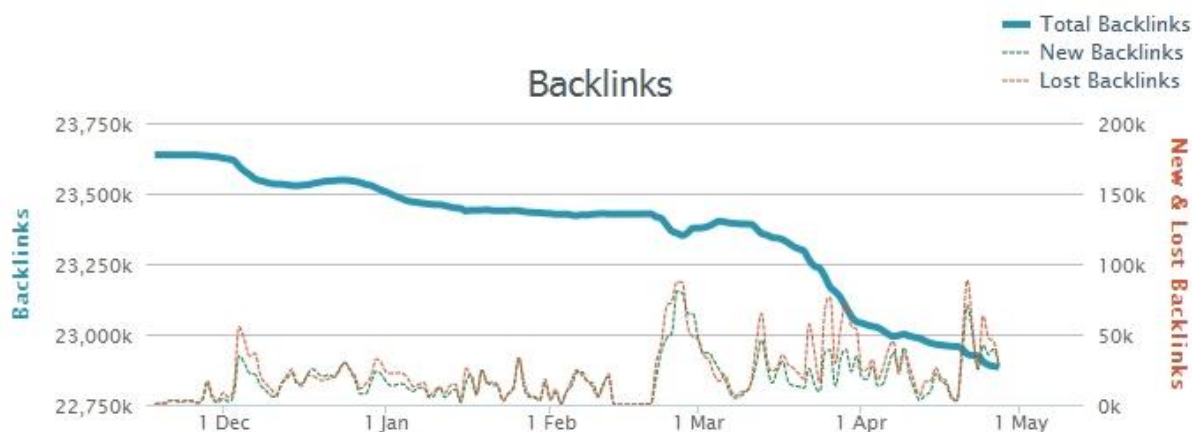
Obr. č. 8: Rýchlosť rastu počtu odkazov na web lenovo.com

(Zdroj: <https://ahrefs.com>)



Obr. č. 9: Rýchlosť rastu počtu odkazov na web pcworld.com

(Zdroj: <https://ahrefs.com>)



Obr. č. 10: Rýchlosť rastu počtu odkazov na web hp.com

(Zdroj: <https://ahrefs.com>)

| STRÁNKA | POČET UŽÍVATEĽOV, KTORÍ MAJÚ STRÁNKU SPOLOČNOSTI VO SVOJICH KRUHOCH (CIRCLES) NA GOOGLE+ <sup>24</sup> | OKNO SO SPRÁVAMI GOOGLE (GOOGLE NEWS BOX) | VÝSLEDKY Z OBLASTI NAKUPOVANIA (SHOPPING RESULTS) | VÝSKYT MEDZI PRVÝMI 20 OBRAZOVÝMI VÝSLEDKAMI NA "LAPTOPS" (IMAGE RESULTS) |
|---------|--|---|---|---|
| A       | plus.google.com/u/0/+BestBuy 5622  | Obr. č. 11                                | objavuje sa                                       | 2 výskyty   |
|         | 5  |   |   | 1   |
| B       | plus.google.com/u/0/+Dell 312 497  | Obr. č. 12                                | objavuje sa                                       | 1 výskyt  |
|         | 2  |   |   | 2   |
| C       | plus.google.com/u/0/+Lenovo 40 154   | Obr. č. 13                                | objavuje sa                                       | 0 výskytov  |
|         | 3  |   |   | 3   |
| D       | plus.google.com/u/0/+PCWorld 5 743   | Obr. č. 14                                | objavuje sa                                       | 2 výskyty   |
|         | 4  |   |   | 1   |
| E       | plus.google.com/u/0/+HP 36 8793  | Obr. č. 15                                | objavuje sa                                       | 0 výskytov  |
|         | 1  |   |   | 3   |

Tabuľka č. 9: Špeciálne pravidlá algoritmu

#### News for **best buy**



Hot Hardware

[Samsung. Best Buy will operate minishops for at least 3 years](#)

CNET (blog) - by Shara Tibken - 21 hours ago

The companies also may introduce Samsung Experience shops to Best Buy stores in Canada. Read this article by Shara Tibken on CNET ...

[Conn's Beats Out Other Big Box Retailers Like Best Buy](#)

Forbes - 23 hours ago

[Best Buy CEO Compensation Valued at \\$19.6 Million in Fiscal 2013](#)

Fox Business - 2 days ago

Obr. č. 11: Zobrazenie bestbuy.com v Google správach

(Zdroj: <https://news.google.com>)

#### News for **dell**

[Is Apple the Next Dell?](#)

Forbes - 5 hours ago

NEW YORK, NY - OCTOBER 06: A copy of the Wall Street Journal and an apple are left in remembrance to Steve Jobs, founder and former ...

[Interview: Dell software chief talks transformation](#)

PCWorld - by James Niccolai - 1 hour ago

[Flawed Bidding Process Leaves Dell at a Loss](#)

New York Times - 1 day ago

<sup>24</sup> Google dáva vyššie hodnotenie autorom a stránkam, ktoré si užívatelia pridali do svojich kruhov



Obr. č. 12: Zobrazenie dell.com v Google správach

(Zdroj: <https://news.google.com>)

#### News for **lenovo**



[China Wireless Targets \*\*Lenovo\*\*, Samsung in China Smartphone Sales](#)

**Bloomberg** - 17 hours ago

China Wireless Technologies Ltd., the nation's third-largest smartphone vendor, said it will eventually overtake market leaders Samsung ...

[Lenovo ThinkCentre Edge 92z](#)

**PC Magazine** - by Joel Santo Domingo - 21 hours ago

[Lenovo ThinkPad Tablet 2 hands on video review](#)

**PC Advisor** - by Chris Martin - 1 hour ago

Obr. č. 13: Zobrazenie lenovo.com v Google správach

(Zdroj: <https://news.google.com>)

#### News for **pcworld**



[LinkedIn revamps contacts feature to boost relationship management](#)

**PCWorld** - 4 hours ago

LinkedIn has revamped the contacts management feature of its professional networking website and packaged the functionality into a new iOS ...

[Windows 8 tablet sales: Actually, not that terrible so far](#)

**PCWorld** - by Jared Newman - 6 minutes ago

[Samsung delays launch of Knox Android security software until 'later date'](#)

**PCWorld** - by Daniel Ionescu - 3 hours ago

Obr. č. 14: Zobrazenie pcworld.com v Google správach

(Zdroj: <https://news.google.com>)

#### News for **hp**

[HP Pavilion TouchSmart 15z-b000 Sleekbook](#)

**PC Magazine** - 3 hours ago

Despite niceties like a large touch-capacitive display and a spacious hard drive, the HP Pavilion TouchSmart 15z-b000 Sleekbook's so-so ...

[HP launches Officejet Pro Black & White all-in-one inkjet printers](#)

**NDTV** - 1 day ago

[HP's first Chromebook goes on sale for £249](#)

**Inquirer** - 1 day ago

Obr. č. 15: Zobrazenie hp.com v Google správach

(Zdroj: <https://news.google.com>)



| STRÁNKA  | POČET TWEETOV | POČET FACEBOOKOVÝCH LIKEOV | ZDIEĽANIA NA FACEBOOKU (SHARE) | POČET PINOV PINTERESTU (PINTEREST PINS) | POČET HODNOTENÍ +1 | AUTORSTVO OVERENÉ PROSTREDNÍCTVOM GOOGLE+ (GOOGLE+ AUTHORSHIP) |
|----------|---------------|----------------------------|--------------------------------|---|--------------------|--|
| <b>A</b> | <b>19</b>     | <b>4 497</b>               | <b>12 592</b>                  | <b>54</b>                               | <b>16 047</b>      | <b>nepoužíva</b>   |
|          | 1             | 1                          | 1                              | 4                                       | 4                  |  |
| <b>B</b> | <b>15</b>     | <b>1 743</b>               | <b>3 864</b>                   | <b>205</b>                              | <b>325 452</b>     | <b>nepoužíva</b>   |
|          | 2             | 2                          | 2                              | 2                                       | 2                  |  |
| <b>C</b> | <b>0</b>      | <b>57</b>                  | <b>163</b>                     | <b>0</b>                                | <b>53 327</b>      | <b>používa<br/>(plus.google.com/+Lenovo)</b>                   |
|          | 4             | 5                          | 5                              | 5                                       | 3                  |  |
| <b>D</b> | <b>0</b>      | <b>425</b>                 | <b>1 570</b>                   | <b>62</b>                               | <b>7 589</b>       | <b>nepoužíva</b>   |
|          | 4             | 4                          | 4                              | 3                                       | 5                  |  |
| <b>E</b> | <b>12</b>     | <b>1 290</b>               | <b>1 750</b>                   | <b>263</b>                              | <b>394 679</b>     | <b>nepoužíva</b>   |
|          | 3             | 3                          | 3                              | 1                                       | 1                  |  |

Tabuľka č. 10: Sociálne signály

| STRÁNKA  | VYHLÁDÁVANIE NÁZVU ZNAČKY A SÚVISIACICH VÝRAZOV - POČET HĽADANÍ MESAČNE (PRIEMER ZA POSLEDNÝCH 12 MESIACOV) | EXISTENCIA VLASTNEJ STRÁNKY ZNAČKY NA FACEBOOK U A JEJ POČET LIKES | EXISTENCIA VLASTNEJ STRÁNKY ZNAČKY NA TWITTER A JEJ POČET FOLLOWERS | OFICIÁLNA STRÁNKA SPOLOČNOSTI NA LINKEDIN | ZOZNAM ZAMESTNANCOV V UVEDENÝ NA LINKEDIN | FYZICKÁ LOKALIZÁCIA POUŽITÍM ZOZNAMU MIEST GOOGLE+ |
|----------|---|--|---|---|---|--|
| <b>A</b> | <b>20 400 000</b> pre výraz "best buy"  | facebook.com/bestbuy<br>6 741 381 likeov                           | twitter.com/twelpforce<br>47 879 Followerov                         | linkedin.com/company/best-buy             | <b>30 763</b> zamestnancov                | <b>používa</b>                                     |
|          | 2   | 1  | 4   |   | 3   |  |
| <b>B</b> | <b>45 500 000</b> pre výraz "dell"  | facebook.com/Dell<br>5 417 902 Likeov                              | twitter.com/Dell<br>66 218 Followerov                               | linkedin.com/company/dell                 | <b>89 919</b> zamestnancov                | <b>používa</b>                                     |
|          | 1   | 2  | 3   |   | 2   |  |
| <b>C</b> | <b>13 600 000</b> pre výraz "lenovo"  | facebook.com/lenovo<br>543 043 Likeov                              | twitter.com/lenovo<br>34 745 Followerov                             | linkedin.com/company/lenovo               | <b>12 806</b> zamestnancov                | <b>používa</b>                                     |
|          | 3   | 4  | 5   |   | 4   |  |
| <b>D</b> | <b>3 350 000</b> pre výraz "pc world"   | facebook.com/PCWorld<br>83 966 Likeov                              | twitter.com/pcworld<br>156 093 Followerov                           | nemá oficiálnu stránku celej spoločnosti  | <b>nemá</b>                               | <b>používa</b>                                     |
|          | 4   | 5  | 1   |   | 5   |  |
| <b>E</b> | <b>2 240 000</b> pre výraz "hp"   | facebook.com/HP<br>2 314 287 likeov                                | twitter.com/HPNews<br>92 256 Followerov                             | linkedin.com/company/hewlett-packard      | <b>187 097</b> zamestnancov               | <b>používa</b>                                     |
|          | 5   | 3  | 2   |   | 1   |  |

Tabuľka č. 11: Signály týkajúce sa značky